ACCURACY OF THE *SPOT*, A PHOTOSCREENING DEVICE, FOR DETECTING REFRACTIVE AMBLYOGENIC RISK FACTORS IN PRESCHOOLERS

A thesis presented to the graduate faculty of New England College of Optometry in partial fulfillment of the requirements for the degree of Master of Science

Hilary L. Gaiser

March, 2016

©Hilary Gaiser, all rights reserved

The author hereby grants New England College of Optometry permission to reproduce and to distribute publicly paper and electronic copies of the thesis document in whole or in part

ACCURACY OF THE *SPOT*, A PHOTO-SCREENING DEVICE FOR DETECTING REFRACTIVE AMBLYOGENIC RISK FACTORS IN PRESCHOOLERS

Hilary Gaiser

This manuscript has been read and accepted by the Thesis Examination Committee in satisfaction of the thesis requirement for the degree of Master of Science.

3/18/2016
Date

Bruce Moore, O.D.
Graduate Faculty Advisor

Catherine Johnson, O.D.
Thesis Committee Member

Nicole Quinn, O.D.
Thesis Committee Member

Jean Ramsey, M.D.
Thesis Committee Member

5/09/2016
Date

Nancy J. Coletta, O.D., Ph.D.
Director of Graduate Studies

New England College of Optometry

Abstract

ACCURACY OF THE *SPOT*, A PHOTOSCREENING DEVICE, FOR DETECTING REFRACTIVE AMBLYOGENIC RISK FACTORS IN PRESCHOOLERS

Hilary L. Gaiser

New England College of Optometry, 2015

Introduction

One of the most common causes of vision disorders in children is refractive error. Refractive errors with the potential to be amblyogenic are of the highest concern, since amblyopia is the most common cause of unilateral vision impairment in people < 60 years of age (Klein, R, 1995). Due to constraints on resources, vision-screening devices such as the Spot, are often used to detect these disorders. The purpose of this thesis is to determine the accuracy of the Spot for the detection of refractive amblyogenic risk factors in preschoolers.

Methods

483 subjects, aged 20 to 99 mos., were recruited from various head starts and early education centers in the Boston area. Subjects were screened by the Spot and then received a masked gold standard examination (GSE). Parental informed consent was obtained prior to vision screening and adhered to the tenets of the declaration of Helsinki. Results of cycloplegic retinoscopy were compared to the non-cycloplegic results of the Spot and were analyzed by using sensitivity/specificity, Bland-Altman and receiver operating characteristic (ROC).

Results

Utilizing the preset referral criteria for the Spot, version 1.1.50, 108 (22.3%) subjects failed and 375 (77.6%) passed the screening. The most common referral reasons were astigmatism 18.63%, myopia 4.35%, anisometropia 2.28%, gaze 0.83% and anisocoria/hyperopia both at 0.41%. Additionally, there were 27 (5.5%) cases of amblyopia, confirmed by GSE using the Vision in Preschoolers (VIP) criteria.

Bland-Altman analysis was used to examine the refractive outputs of the Spot compared to cycloplegic retinoscopy for sphere, cylinder and anisometropia. For spherical error the bias was 1.25 D OD and 1.39 D OS. The 95% limits of agreement for the right eye was upper 3.24 D and lower -0.74 D and left upper 1.40 D and lower -0.65 D. There was a statistically significant, positive proportional bias (trend) for both the OD and OS. As the average of the two measures increases by 1 D, the bias increases by 0.51 D ($P< 0.00$) diopters for the OD and 0.57 ($P<0.00$) D OS. For cylinder, the bias was 0.37 D OD and 0.27 D OS. The 95% limits of agreement for the OD was upper 1.49 D and lower -0.75 D and OS upper 1.40 D and lower -0.87. There was a significant positive proportional bias (trend) for the OS but not OD. As the average of the two measures increases by one diopter, the bias increases by 0.03 ($P<0.39$) D OD and 0.13 ($P< 0.00$) D OS. For spherical anisometropia, the bias was found to be -0.03 D. The 95% limits of agreement were upper 0.81 D and lower -0.87 D. There was also a statistically significant, proportional bias (trend). As the average of the two measures increases by one diopter, the bias increases by 0.14 D ($P< 0.00$) D.

With the Spot's pre-set criteria software (version 1.1.50) the Spot's sensitivity for detection of amblyopia was 77.78% CI (57.26%-90.62%) and had a specificity of 80.44% CI (76.43%-83.92%). When the VIP criteria was applied to the refractive error output of the

Spot, the sensitivity of the Spot for amblyopia detection was 77.77% CI (57.27, 90.62) and the specificity was 77.70% CI (73.53, 81.39). Using the modified referral criteria that was determined using (ROC) analysis the sensitivity of the Spot was 85.18% CI (65.39, 95.14) and the specificity was 77.04% CI (72.84, 80.78). However, there was a mismatch between the reason for referral of the Spot (preset criteria) and the cause of amblyopia as determined by GSE.

Using ROC analysis for each eye separately, the area under the curve (AUC) for hyperopia (defined as wet sph >/= +3.25 D) was 0.86 (0.80, 0.93) OD and 0.86 (0.81, 0.92) OS. Astigmatism (defined as wet exam >/= 1.50D) AUC was found to be 0.93 (0.90, 0.96) OD and 0.91 (0.87, 0.95) OS. When analyzing the eye with the highest refractive error of each subject the AUC was as follows, most hyperopia meridian (0.74), cylinder power (0.85), anisometropia sph. (0.66), anisometropia cyl. (0.66) and with all four predictors combined (0.88). By using ROC when specificity for the detection of amblyogenic risk factors (defined as hyperopia >/= +3.25, astigmatism >-1.50D, anisometropia sphere > 1D, cylinder > 1D) was set at 90% the optimal pass/fail criteria of the Spot was determined. The largest magnitude criteria setting that would maximize sensitivity is, hyperopia Sph 1.75 D, Cyl -1.50 D, anisometropia Sph 1.25 D and anisometropia Cyl 1.25 D. This results in a sensitivity of 73% for the detection of amblyogenic risk factors.

Discussion

In conclusion, the Spot appears to be accurate in the detection of lower level hyperopia (defined as < +3.25 D) and astigmatism. The Spot has difficulty detecting higher levels of hyperopia, > +3.25 D, which is consistent with other studies. This may be due to

inherent problems with the photoscreening technology of the device and not due to a lack of accommodation control. The average pupil size was similar in both those subjects < +3.25D and > +3.25 D which would indicate that accommodation may have been adequately controlled in both groups. Additionally, the Spot's preset referral criteria v.1.1.50 had a much lower referral rate for hyperopia (0.41%) than expected for this population. This may be due in part to the use of spherical equivalent, which would underestimate hyperopia in a high astigmatic population.

For the detection of amblyopia, the sensitivity and specificity of the Spot appears to be good and is consistent with the findings in other studies. However, specificity and sensitivity do not take into account the Spot's reason for referral compared to GSE and it may be artificially high depending on the study population.

In order to increase the accuracy of the Spot device, the optimal referral criteria was determined by ROC analysis. This optimal referral criteria increased the Spot's sensitivity for the detection of amblyopia and increased the agreement in the reason for referral between the Spot and GSE. The new referral criteria is significantly lower than the current preset criteria and is similar to other optimal referral criteria proposed in other studies.

Acknowledgements

First and foremost I would like to thank my mentor Dr. Bruce Moore for his guidance on this project and constant encouragement. I also want to thank him for introducing me to the world of pediatric optometry and vision screening which has greatly influenced my optometric career.

Additionally, I would like to thank Catherine Johnson OD, Nicole Quinn OD and Jean Ramsey MD for serving on my committee and for their feedback on this project. This project would also not be possible if it wasn't for the support of NEEI mobile eye clinic particularly Paulette Tattersall, Ed Braverman, Farida Layakoubi and Gary Chu OD. I would also like to thank Dr. Diane Ah-Kine OD for her help in providing the gold standard exams.

I want to give a big thank you to Dr. Nancy Coletta and the department of graduate studies for an amazing experience as a dual OD/MS in vision science student. I would also like to thank Dr. Li Deng for her bio-statistics instruction, data analysis help and for answering all my statistics questions.

I would also like to thank my fellow MS students, Nadine Solaka, Ran He and Blair Germain for their assistance on this project as well. Last but not least I would like to thank my wonderfully supportive boyfriend, Kevin Rogers, for putting up with me during this process and for helping me with my emergency IT and statistics questions. I give a big and hearty thank you to you all.

Table of Contents

Abstract……iii

Acknowledgements ……vii

Table of Contents……viii

List of Figures and Tables ……x

Chapter 1: Introduction ……..…..1
1.1 Amblyopia……..….1
1.2 The Pathogenesis of Amblyopia……...……..4
1.3 The Impact of Amblyopia…….....….5
1.4 Prevalence of Amblyopia…….....6
1.5 Impact of Uncorrected Refractive Error……..…….....7

Chapter 2: Vision Screening…… ……...8
2.1 Eye Examination Rates in Children……..……..8
2.2 World Health Organization (WHO) Screening Guidelines……...…….9
2.3 Vision Screening Recommendations in the United States…….....9
2.4 Nationwide Screening Laws in the United States…….....12
2.5 Vision Screening Rates in the United States…….....13
2.6 Vision Screening Personnel……..…….....14

Chapter 3:Vision Screening Devices and Device Validation……14
3.1 Vision Screening Device Regulations…….……..……14
3.2 Validation of Vision Screening Devices……...…….16
3.3 Vision Screening Devices……..……....17
3.4 Spot……..……..18
3.5 Summary of Spot Studies……..22

Chapter 4: Methods and Objectives……..…….23
4.1 Purpose, aims and objectives…...……..23
4.2 Methods……..24
4.2.1 Subject Recruitment and Data Collection……..…….……...24
4.2.2 Screening Protocol……..25
4.2.3 Gold Standard Examination……..……..……26
4.3 Statistical Methods…….28
4.3.1 Sensitivity and Specificity……..…….28
4.3.2 Receiver Operator Characteristic (ROC) analysis……..……29
4.3.3 Bland-Altman Analysis ……..……29

Chapter 5: Results……..……30
5.1 Subject Characteristics and Data Overview……..……..……..30
5.2 Measurement Agreement of Spot and GSE…….....32
5.2.1 Bland-Altman Spherical Analysis……..……33
5.2.2 Bland-Altman Cylindrical Analysis……...…34
5.2.3 Bland-Altman Spherical Anisometropia……...…….36
5.3 High Hyperopic Subjects and Average Pupil Size Measurement…….……36
5.4 Sensitivity and Specificity for the Detection of Amblyopia…….……38
5.3.1 Spot Criteria Software v.1.1.50……..……39
5.3.2 VIP Criteria……..……...40
5.3.3 Modified Criteria……..……..42
5.5 ROC analysis…….43
5.4.1 ROC for Right and Left Eyes ……..44
5.4.2 ROC for Eye with Highest Refractive Error (HRE)……..……45
5.4.3 Modified Referral Criteria……..46

Chapter 6:
Discussion……...46
6.1 Data overview…….……46
6.2 Spherical Measurement agreement…….……...50
6.3 Cylindrical Measurement Agreement…….…51
6.4 Spherical Anisometropia Measurement Agreement……..……...…..52
6.5 Sensitivity and Specificity for the Detection of Amblyopia……...53
6.6 ROC analysis for Right and Left Eyes……56
6.7 ROC analysis for the eye with highest refractive error……57
6.8 Recommended Optimal Referral Criteria Guidelines for Spot……....59
6.9 Pupil Size and Accommodation……..…….....……..59
6.10 Study limitations ……61
6.11 Summary and further work……62

Appendices

Appendix I: Summary of AAPOS State-by-State Vision Screening requirements…...……64

Appendix II: Summary of Spot Studies…….…….…….....67

Appendix III: Cutoffs for sensitivity and specificity with specificity set at 90%.....69

Bibliography……...……70

List of Figures and Tables

Figure 1: Reason for referral of Spot's pre-set criteria…….....31

Figure 2: Bland-Altman Sph. Right Eye……..33

Figure 3: Bland-Altman Sph. Left Eye……..……..34

Figure 4: Bland-Altman Cyl. Right Eye……..35

Figure 5: Bland-Altman Cyl. Left Eye……35

Figure 6: Bland-Altman Spherical Anisometropia……..……...36

Figure 7: OD Difference in measurement <>+3.50 D……...…..……37

Figure 8: OS Difference in Measurement <> +3.50 D……38

Figure 9: ROC Hyperopia OD and OS……44

Figure 10: ROC for Astigmatism OD and OS…….……45

Figure 11: ROC Curves for Comparison EHRE……...……..……45

Table 1: Summary of Refractive Amblyogenic Risk Factors……..……..2

Table 2: Pre-Set Referral Criteria for Spot v.1.1.50……...…...…..…21

Table 3: Measurement Range and Reported Accuracy of Spot…….....…….……...22

Table 4: Contingency table Spot Criteria v.1.1.50…….…….……..……...39

Table 5: Reason for referral in subjects with amblyopia (preset criteria)……40

Table 6: VIP Amblyogenic Risk Factors……..……..40

Table 7: Contingency table for VIP referral criteria……...……..41

Table 8: Reason for referral in subjects with amblyopia (VIP criteria)…….……..41

Table 9: Contingency table based on modified referral criteria……..……42

Table 10: Reason for referral in subjects with amblyopia (modified criteria) ……..……..43

Table 11: Summary of Referral Reason for Spot……...….48

Table 12: Comparison of device AUC……..……..58

Image 1: Patient view of Spot…….……..……19

Image 2: Spot in use…….20

Chapter 1: Introduction

Some of the most common vision disorders in children are caused by refractive error. Refractive errors with the potential to be amblyogenic are of the highest concern, since amblyopia is the most common cause of unilateral vision impairment in people < 60 years of age (Klein, 1995).

1.1 Amblyopia

Classically, amblyopia has been defined as a loss of best-corrected visual acuity in one or both eyes despite the absence of disease. It is considered a cortical deficit that occurs secondary to high isoametropic or anisometropic refractive error, strabismus, occlusion of the visual pathway or a combination of any of the preceding categories. (Steinman, S. 2000). Additionally, amblyopic eyes can have deficits not measured by visual acuity alone such as problems with accommodation, fixation and tracking (Bedell, HE. 1985).

Refractive Amblyopia

Amblyopia resulting from high refractive errors is the most common. High hyperopia, astigmatism and anisometropia are the typical refractive errors associated with amblyopia, although rarely myopia can also be at fault. The magnitude of refractive errors deemed amblyogenic is currently debated and many eye care professional organizations have released guidelines. A summary of current guidelines is found in Table 1.

Table 1: Summary of Refractive Amblyogenic Risk Factors

American Optometric Association (AOA 2004)				
ISO.	Hyperopia	Myopia	Astig.	
	>5D	>8D	>2.50	
Aniso.	>1 D	>3D	>1.5D	
Vision in Preschoolers Study (VIP 2004)				
	Hyperopia	Myopia	Astig.	
ISO.	>3.25D	>2 D	>1.50	
Aniso.	>1D	>3D	>1.5D	
Antimetropia	One eye >1D hyper and >1D between eyes	One eye >2D myopia and >3 D between eyes		
American Academy of Ophthalmology* (AAO 2012)				
ISO.	Hyperopia (no deviation)	Myopia	Astig.	Hyperopia + esotropia
<1 year	>6D	>5D	>3D	>2.5D
1-2 years	>5D	>4D	>2.5D	>2D
2-3 years	>4.5D	>3D	>2D	>1.5D
Aniso. (no strab)	Hyperopia	Myopia	Astig.	
<1 year	>2.5D	>4D	>2.5	
1-2 years	>2D	>3D	>2D	
2-3 years	>1.5D	>3D	>2D	
American Association for Pediatric Ophthalmology and Strabismus (Donahue 2013)				
	Hyperopia	Myopia	Astig.	Aniso.
12-30 mos.	>4.5D	NA	>2.0D	>2.5D
31-48 mos.	>4D	NA	>2D	>2D
>49 mos.	>3.5D	NA	>1.5D	>1.5D

AAO *specific guidelines for older children are not provided as correction is determined by severity of refractive error, VA and patient symptoms

Strabismic Amblyopia

Strabismus is usually categorized by age of onset (infantile or acquired), type of deviation (exotropia, esotropia), frequency of deviation (intermittent versus constant) and location (left, right or alternating). Strabismus typically causes amblyopia when the deviation occurs before the critical period of development and is constant and unilateral in nature. The deviating eye will have abnormal visual input compared to the non-deviating eye and as a result of this abnormal vision the deviating eye will tend to suppress the central visual field. The suppression phenomenon is theorized to be an adaptation to help the better seeing eye (Von Noorden and Campos, 2002).

Deprivation Amblyopia

Deprivation amblyopia is considered the most serious type of amblyopia in terms of visual acuity. It is also the type of amblyopia most closely connected with some type of ocular pathology. Typically, it is caused by a constant obstruction of the visual pathway that includes but is not limited to cataract, ptosis, corneal opacities and even excessive or improper patching that occurs during the critical period of development. Even brief periods of deprivation during the early stages of development can cause amblyopia since a normal retinal image is never formed and subsequently proper visual development cannot occur (Steinman, 2000). While deprivation amblyopia is the rarest type of amblyopia <3% of cases, it is the most likely to cause severely reduced vision and necessitates the most aggressive treatment. (Hillis, 1983).

1.2 The Pathogenesis of Amblyopia

Unilateral amblyopia is caused by suppression of the image of the most blurred eye and can occur in cases of anisometropic, strabismic and deprivation amblyopia. This suppression of the blurred eye allows for a clearer image in the less blurred eye, which is preferred by the visual system. In amblyopia, the poorer seeing eye is suppressed to the point that the vision remains abnormal even when the clearer eye is covered and/or the blurred eye is fully corrected. (Steinman, 2000). It should be noted that unilateral amblyopia is not purely a disease of only one eye as the tracking on the non-amblyopic eye can also be less accurate than in age matched controls (Bedell, 1985).

The physiology of the suppression phenomenon has been debated but may be due in part to binocular rivalry and/or dichoptic masking. In normal binocular rivalry, the observer experiences alternating eye dominance and suppression of each monocular image when corresponding points in the two eyes are so dissimilar that they cannot be fused. This mechanism may occur as an attempt by the visual system to prevent diplopia. In amblyopia, binocular rivalry is more extreme and may be one of the causes of the suppression of the central visual field. Dichoptic masking is the process where a stimulus of one contrast presented to one eye can prevent the detection of a lower contrast image in the other eye, despite having an otherwise identical stimulus. This binocular inhibitory process can be found in some amblyopic individuals and may account for the suppression found in some small angle strabismus (Harrad, 1996).

In the case of bilateral or isoametropic amblyopia, both eyes have similar refractive errors but the magnitude of the refractive error is to such a large degree that neither eye

receives a clear, focused image, resulting in decreased acuity for both eyes due to poor brain-eye connections. This is particularly true in the case of high hyperopia where the patient cannot accommodate enough to compensate for the refractive error. In the case of myopia, there is a defined point at which there is clear vision, which makes it less likely to cause amblyopia. Since the two eyes are similar, this type of amblyopia is typically less dense amblyopia than that found in unilateral cases (Steinman, 2000).

Meridional amblyopia, resulting from high astigmatism, is typically the mildest form of amblyopia. In these cases, one meridian is usually relatively clear while the other is blurred. This can cause reduced acuity, again due to poor eye-brain connections during the critical period, along one meridian even when fully corrected. This typically presents as reduced visual acuity for particular letters that fall on the most blurred meridian (Steinman, 2000).

1.3 Impact of Amblyopia

Currently, the cost-effectiveness and psychological impact of amblyopia treatment is being debated (Membrano, 2002 and Rein, 2012). However, studies indicate that there is a significant visual impact on an individual with amblyopia. The most striking visual impact is associated with ocular injury to the better seeing eye resulting in poorer overall vision. Van Leeuwen et al. (2007) concluded that the estimated lifetime risk of bilateral vision impairment nearly doubled in those with amblyopia.

Additionally, in an Australian study, the consequences of amblyopia on social class were examined. It was noted that although amblyopia was not significantly associated with

lifetime occupational class, people with amblyopia ultimately completed fewer higher university degrees. The same study also found that people with amblyopia had almost three times the risk of visual impairment (vision < 6/12 or < 20/40) in the better eye compared to people without amblyopia (Chua, B. 2004). Yet another study from the UK noted that the projected lifetime risk of vision loss for an individual with amblyopia is 1.2%. Of those amblyopic individuals with additional vision loss in the better eye only 35% of people in previously paid employment were able to continue (Rahi, 2013).

Other detrimental effects on vision include problems with spatial-temporal vision, reading speed and lack of binocular vision. In general reduced stereopsis can cause problems with activities requiring spatial certainty, can affect control of manipulation, reaching and balance (Jones, RK. 1981 and Joy S. 2001) and difficulty with tasks that rely on three-dimensional clues (Murdoch, 1991). This lack of binocularity, can potentially impact an individuals ability to participate in certain professions and activities.

1.4 Prevalence of Amblyopia

The prevalence of amblyopia varies across studies. This variation can be attributed to differences in the definition of amblyopia and the populations studied. In the literature, the range of prevalence of amblyopia varies from 0.5% to 5.3% but around 3% is the most common (Polling, JR. 2012, Jamali, P. 2009 and Von Nordon, 2002).

1.5 Impact of Uncorrected Refractive Error on Education

Previously the impact of uncorrected refractive error, particularly hyperopia, on the education of preschoolers has been somewhat anecdotal. Studies examining this topic have been small and have had conflicting results (Lamoureux, 2009).

Recently, the results of a larger study by the VIP-HIP group examining uncorrected hyperopia and preschool early literacy has been released. In this cross sectional study of children attending preschool or kindergarten who had not previously worn refractive correction, the early literacy of 4 and 5 year old uncorrected hyperopic children compared with emmetropic children was examined. Out of a total of 492 children (244 hyperopes and 248 emmetropes) the mean difference in the Test of Preschool Early Literacy (TOPEL) scores between the two groups was found to be -4.3 (P=0.01) for TOPEL overall, -2.4 (P=0.007) for print knowledge, -1.6 (P=0.07) for definitional vocabulary and -0.3 (P=0.39) for Phonological Awareness. There was even greater differences for hyperopic individuals >/= + 4 D. These were found to be −6.8 ($P = 0.01$) for total score and −4.0 ($P = 0.003$) for Print Knowledge. The greatest differences between groups was for hyperopic subjects with a binocular near VA of </= 20/40 or stereoacuity of </= 240 sec of arc. These children were found to have scores of −8.6 ($P < 0.001$) for total score and −5.3 ($P < 0.001$) for Print Knowledge (VIP-HIP, 2016).

Chapter 2: Vision Screening

2.1 Eye Examination Rates of Children

While efficacy and cost-effectiveness of pediatric vision screening is debated (Powell, 2009), rates of eye examination in children, excluding screening, remain low. A 2002 report by the Vision Service Plan noted that by age 12, only 52% of children have been seen by an eye care professional. Typically parents only sought eye care for older children, instead of at younger ages when amblyopia treatment would be most effective (Martinez, 2002).

Another aspect to consider when trying to increase the rates of eye examination is the number of available providers. According to the United States Bureau of Labor Statistics there are an estimated 33,340 practicing Optometrists in the United States with an additional 1,125 added each year (US Dept. Labor, 2014). Additionally, as of 2012 there were an estimated 18,805 practicing Ophthalmologists in the US (ICO, 2015). This makes an estimated total of 52,145 eye care professionals. In comparison, US Census Data indicates that as of 2013 there were an estimated 73,616,221 children in the United States under the age of 18. This means that each provider would have to see ~ 1412 kids yearly, ~ 27 weekly, if every child were to have an annual exam. (USCB, 2015). Theoretically, the number of eye care providers should be sufficient to handle the demand but due to the aging US population, location of providers and variability of provider's ability to provide pediatric services there is likely still unmet need for vision services. Highly specific/sensitive vision screening offers a solution for this unmet need as those with the highest need can be identified and referred for the appropriate services.

2.2 World Health Organization Screening Guidelines

According to the WHO guidelines initially published in 1968, in order for a screening program to be effective it should adhere to the following guidelines: the condition should be an important health problem, there is a treatment for the condition, a latent stage of the disease exists and a test or examination that is acceptable to the population is available for detection. Additionally, the natural history of the disease should be adequately understood and a standard policy on whom to treat and the total cost of finding a case should be economically balanced in relation to medical expenditure as a whole (Wilson JMG, 1968).

Based on these criteria, screening for amblyopia adheres to the WHO guidelines. Amblyopia is an important health problem that impacts an individual's life. While the most accurate and efficient screening tests for amblyopia are currently debated, the tests typically involved are non-invasive and are generally accepted by the population. The prognosis and/or natural history of amblyopia as well as its treatment has been extensively studied and it is generally agreed upon that early treatment promotes the best outcome. Additionally, amblyopia treatment has been determined to be cost-effective (Membrano, 2002).

2.3 Vision Screening Recommendations in the United States

Currently, the United States Preventative Services Task Force recommends vision screening for all children at least once between the ages of 3 and 5 years in order to detect the presence of amblyopia or its risk factors (grade B recommendation). For those children 0 to 3 there is currently insufficient evidence to issue a recommendation for or against vision screening. This is mainly due to a lack of evidence that early detection leads to better

outcomes and currently there are no evidence based techniques to screen this age group effectively (USPSTF, 2011). A research project is ongoing at the time of this writing to develop an efficient screening protocol for infants and toddlers. This project uses a questionnaire given to parents to assess abnormal visual behavior along with screening by the Spot compared to gold standard examination.

For those children over the age of three, an agreed upon battery of screening tests has not been fully established but recommendations have been released by both a National Expert Panel and the American Academy of Pediatrics. The best practice guidelines released by the National Expert Panel outline testing procedures for children between the ages of 36 to 72 months. Two main vision screening methods are recommended, monocular vision acuity testing and instrument based testing using autorefraction but not photoscreeners (Cotter, 2015).

When using monocular visual acuity tests, tests that are made up of single crowded HOTV letters or Lea Symbols surrounded by logMAR scaled crowding bars should be used. According to the most recent best practices, these charts should be calibrated to a 5 ft (1.5m). However, older tests used at a testing distance of 10 ft are also acceptable. The child can either verbally call out or match the optotypes. Other optotypes such as Allen symbols or Tumbling E should not be used.

In terms of visual acuity level used as a cut-offs for referrals, a passing score is correctly calling 3/3 or 3/4 with each eye at the 20/50 level for ages 36 to 47 months (ages 3 to ~ 4) and at the 20/40 level for aged 48 to younger than 72 months (ages 4 to 6). This is for the newer single surrounded HOTV letters/Lea symbols calibrated for 5 ft. Older tests, such

as the rectangular crowding bar surrounding a single line of HOTV letters/Lea symblols calibrated for 10 ft, utilizes a slightly different referral criteria, and is now considered acceptable but not best practice. Using the PASS stereoacuity test as an adjunct to supplement either one of the best practice models is also acceptable. (Cotter, 2015).

If device based screening is chosen, then an autorefractor such as the Retinomax or SureSight Vision Screener should be used. (Cotter, 2015). Photoscreeners, including the MTI Photoscreener, Power Refractor II and iScreen Photoscreener, had been evaluated by the VIP group and were found to have unsatisfactory sensitivity in detecting amblyopia, strabismus and significant refractive error compared to autorefractors such as the Retinomax and SureSight. However, the Plusoptix photoscreener has been labeled as acceptable but not best practices (VIP, 2004). It should be noted that this recommendation is based on devices used in the VIP study a decade ago and does not necessarily include the most recent devices. The Spot has recently been added as a recommended screening device (Cotter, 2015).

The best practices guidelines also state that screening should be performed annually in this age group. The minimum standard for vision screening should be at least once during the interval between their 3rd and 6th birthdays. Children who should automatically be referred for a complete vision exam include those with neurodevelopmental disorders, systemic conditions associated with ocular abnormalities, first-degree relatives with amblyopia or strabismus, those who were born prior to 32 weeks and any child whose parent continues to express concern about a potential vision problem. Any child who refuses vision screening or who fails to complete vision screening should be considered an automatic referral (Cotter, 2015). The VIP Group noted that children who failed to complete screening

or who were untestable during screening are more likely to have vision issues and should be rescreened or sent for a complete eye exam no later than six months after the initial uncompleted screening (VIP, 2007).

The American Academy of Pediatrics has also issued a policy statement regarding vision screening. While there is no recommendation for mass screening using device-based vision screening at this time. Photoscreening and handheld autorefraction are recommended as an alternative to visual acuity screening from ages 12 months and 3 years at annual well child visits. Direct testing of visual acuity can begin by age 4 using age appropriate targets. As noted by this group instrument based screening may be a helpful alternative in screening developmentally delayed children of any age. For children younger than 12 months, such as infants, red reflex testing is recommended (AAP, 2016).

2.4 Nationwide Screening Laws in the United States

Laws regulating and mandating vision screening in children vary widely and depend upon state laws. Additionally, there is often little guidance for performing vision screening and at which ages screening is most appropriate. This can create a burden for schools and pediatricians, particularly those with limited resources.

According to data compiled by AAPOS in 2011, out of fifty states plus the District of Columbia and Puerto Rico 37 states mandate some type of vision screening while 11 do not. Four states require a complete eye exam and one state, Alabama, requires a general medical exam but does not indicate if vision is included. Illinois both requires a full eye exam prior to entry into school and subsequent vision screenings afterwards and falls into both the vision

screening and vision exam categories. Appendix I contains a summary of vision screening laws.

2.5 Vision Screening Rates in the United States

Even with screening guidelines and legislation in place, it is often difficult to measure how much they are fully implemented into a screening program. Recently a quality improvement project to optimize chart based vision screening for preschool aged children was implemented and studied. The completion of full vision screening (monocular distance VA plus stereovision) was documented at well-child visits in both a pre and post intervention period. The completion of full vision screening was 35.6% of children in the pre-intervention period and 41.3% in the post-intervention period. When stratified by age, screening improved for children aged 5 and 4 but declined among 3 year olds. This study demonstrates that even with good screening protocols in place it is difficult to accomplish high rates of vision screening using chart-based techniques and stereovision in preschoolers (Modest, 2016).

Studies have also shown a wide range of vision screening rates in the US preschool population. These rates vary from 2% to 64% based on the type of screening, population and the ages of children. (Levison, 2013, Hambridge, 2007, Shaw, 2006, Stange, 2000, Marsh-Tootle, 2008). Together these studies demonstrate that easy to use vision screening technology and not visual acuity and stereovision is likely to be the best method to accomplish universal vision screening at well-child visits in young children. (Modest, 2016).

2.6 Vision Screening Personnel

Due to limited resources, lay-screeners such as untrained school nurses, parents and teachers are often employed for vision screening purposes. When using such individuals for vision screening, easy, simple to use screening tests are recommended. According to the Vision in Preschoolers Study Group, automated screening devices such as the Retinomax Autorefractor and SureSight Vision Screener are as effective when used by trained nurse screeners and by lay screeners as when they are used by licensed eye professionals (LEPs) such as optometrists and ophthalmologists. This is in contrast to linear Lea Symbols VA screening which is more effective when used by LEPs than by lay screeners (VIP, 2005). Thus, if lay-screeners are performing vision screening services it is highly likely that device based vision screening will be the most effective means of vision screening.

Chapter 3. Vision Screening Devices and Device Validation

3.1 Vision Screening Device Regulations

In the United States, vision screening devices are considered medical devices and are subject to regulatory controls. Medical devices are classified into Classes I, II and III according to FDA guidelines. Regulatory control increases with the Class number. The Spot is classified as a Class I device, ophthalmic refractometer, and as such is exempt from Premarket Notification 510(k). Premarket Notification 510(k) means that a device cannot be commercially distributed until the FDA has determined that the device is substantially equivalent to another device already in legal commercial distribution in the United States or

to another substantially equivalent device. Approximately, 74% of Class I devices are exempt from this pre-market notification according to the FDA (FDA, 2015).

To be exempt from pre-market notification the FDA states, "the exemption from the requirement of premarket notification…."for the generic type of Class I or Class II device is only to the extent that the device has existing or reasonably foreseeable characteristics of commercially distributed devices within that generic type or, in the case of in vitro diagnostic devices, only to the extent that misdiagnosis as a result of using the device would not be associated with high morbidity or mortality." (FDA, 2015).

FDA classification of a device is based on risk assessment to both the patient and the user. Device classification also depends on the intended use of the device and its indications for use. All classes of devices are also subject to general controls that are the baseline of the Food, Drug and Cosmetic Act (FD&C) Act. (FDA, 2015). General controls apply to all medical devices and exist as a means to ensure their safety and effectiveness. They include provisions that relate to adulteration, misbranding, device registration and listing, premarket notification, banned devices, notification, including repair, replacement, or refund, records and reports, restricted devices, and good manufacturing practices (FDA, 2015).

Before placing a device in Class I, the FDA must determine that the device can be listed by this classification and that the general controls are enough to assure the safety and effectiveness of the device. Class I devices are not intended to: 1) support or sustain life 2) be of importance in preventing impairment to human life 3) present a potential unreasonable risk of illness or injury (FDA, 2015).

3.2 Validation of Vision Screening Devices

In the United States, no intensive device validation studies are currently required for vision screening devices that fall into the Class I 510(k) exempt designation. While these devices may be classified as low risk of harm to the public, there is often little validation of their effectiveness. The WHO has released guidelines on medical device regulatory systems and advocates having medical treatment theories and technological concepts validated by controlled scientific experiments before being released into the general public.

Under the WHO's classifications, device regulatory systems fall into "safety only" systems or safety and demonstration of clinical effectiveness/performance prior to market release. Under a safety only system, such as the FDA Class I 510(k) exempt, if a device is safe and can perform its medical function it would then be allowed market release. Device effectiveness would then be measured by general consensus by health care professionals, medical literature and consumers based on product use. It is assumed that ineffective devices would acquire a poor reputation, demand would diminish and ultimately production would cease (WHO, 2001).

However, the second approach considers device safety inseparable from effectiveness. While assessing effectiveness may cost more in time and money it can intervene to prevent negative outcomes for patients. It takes into account that before consensus can be reached by health care professionals, a significant number of patients may have received ineffective treatments/procedures. Consensus can also be vastly inaccurate as individual health care professionals may treat too few patients under poorly controlled conditions to come up with an accurate conclusion. Additionally, financial resources have

been used to support these ineffective treatments and more financial resources will need to be spent to support those who have received ineffective treatment/procedures (WHO, 2001).

Ideally, all vision-screening devices, would go through rigorous comparison to a masked gold standard examination in order to evaluate their effectiveness for determining significant refractive error. Until a vision screening device is considered equal to a GSE comparing vision screening devices to each other without comparison to GSE lacks scientific rigor.

3.3 Vision Screening Devices

Automated vision screening machines consist of two main groups, autorefractors and photoscreening devices. Both groups screen by identifying the magnitude of refractive error and comparing it to age-specific pass/fail refractive error criteria. In some devices, the operator must interpret the refractive error measurement output as a pass/fail while in others a comparison to pre-set pass/fail criteria is included as part of their software (Cotter, 2015). The types of devices used for vision screening use different optical principles and confusion is created when names are used interchangeably.

Traditional autorefractors typically use infrared technology to measure the size of a reflected image, usually a ring, to determine refractive error. If the reflected light exits in collimation the patient is emmetropic, convergent if myopic and divergent if hyperopic. A more recent type of device uses ocular wavefront technology and is also commonly classified as an autorefractor. This type of technology generally uses the Hartmann-Shack principle. In simple terms, light from a spot on the retina is scattered as it exits the corneal plane, creating

a wavefront that is then captured and analyzed by a camera (Lebow, 2014). In autorefractors, measurement of refractive error is monocular and some method to control accommodation in the non-tested eye is used (Cotter, 2015).

Photoscreening devices generally use an optical image of the eyes' red reflex to provide a simultaneous, binocular estimate of refractive error. As with autorefractors a means of controlling accommodation, which varies among devices, is also employed (Cotter, 2015). The basic principles for photoscreening devices are based on a device by Schaeffel, Farkas & Howland that was developed in 1987. This device consisted of rows of infrared LEDs positioned at different distances, or eccentricities, from the edge of a shield that occluded part of a video camera lens. After light from the LEDs is projected into the eye, a small proportion is also reflected from the fundus back to the camera. If the light is not in focus due to refractive error the reflected light will spread out in a cone. The amount of defocus changes the angle of spreading light, which in turn changes the reflex. It is the position of the reflex that is used to help determine the type of refractive error. For a hyperope, the reflex would only be seen in the upper portion of the eye while for a myope this reflex would only be seen in the lower portion of the pupil. In an eye that is perfectly focused the pupil appears almost dark. In high refractive errors the entire pupil is flooded with light and multiple eccentricities have been used to try to extend the range of measurement (Schaeffel 1993).

3.4 Spot

According to information supplied by Welch Allyn, the Spot device uses invisible infrared lights to pulse a pre-determined pattern to capture the refractive error measurement.

The optical design specific to the Spot includes infrared light sources both inside and outside the lens area in order to eliminate the "dead zone" reported by typical photoscreeners (Welch Allyn, 2015). This dead zone is a result of an insensitive range of refractions detection due to a totally dark reflex. (Shi, 2011). This helps to increase the range of refractions measured and reduces alignment issues. Additionally, the capture rate is adjusted dynamically in order to provide consistent signal levels for all subjects despite eye color (Welch Allyn, 2015).



Image 1: Patient view of Spot

The device is portable and is held approximately 3 feet from the subject for screening purposes. As part of its software, the Spot gives feedback to the screener about the distance from the subject, such as too far or too close in order to facilitate capture. Additionally, the

device emits sounds that help direct the attention of the subject to the device. In order to control accommodation and to maintain attention flashing lights are utilized.

The device also comes with modifiable pre-set pass/fail criteria that flags a subject's measurements as either a pass (green) or fail (red). The preset referral criteria v. 1.1.50 is listed in table 2. Subject data is entered via a touchscreen and includes: Subject ID, first name, last name, date of birth, gender and eyewear status. Eyewear status includes none, glasses or contacts.



Image 2: Spot in use

Table 2: Pre-Set Referral Criteria for Spot v.1.1.50

Age (Mos.)	Hyperopia	Astigmatism	Myopia	Anisometropia	Anisocoria
12-36	3.00 D	-2.00 D	-2.00 D	1.00 D	1 mm
36-72	2.50 D	-1.75 D	-1.25 D	1.00 D	1 mm
72-240	2.50 D	-1.50 D	-2.00 D	1.00 D	1 mm

The measurement range and reported accuracy of the Spot can be found in table 3. Although the Spot provides some degree of measurement of eye alignment, analysis of its accuracy is beyond the scope of this thesis.

Table 3: Measurement Range and Reported Accuracy of Spot

Spherical Equivalent	
Range	-7.50 D to +7.50 D in 0.25D increments
Accuracy	-3.50 D to 3.50 D +/- 0.50 -7.50 D to <-3.50 D +/- 1.00 >3.50 D to 6.00 D +/- 1.00
Cylinder	
Range	0.00 to +3.00 D in 0.25 increments
Accuracy	0.00 to 1.50D +/- 0.50 >1.50 to 3.00 D +/- 1.00D
Axis	
Range	1 to 180 degrees in 1 degree increments
Accuracy	+/- 5 degrees >0.50 D
Pupil Size	
Range	4.0 mm to 9.0 mm in 0.1 mm increments
Accuracy	+/- 0.1 mm

3.5 Summary of current Spot studies

While many studies have attempted to determine the accuracy of the Spot for its ability to detect amblyopia and amblyogenic risk factors, these studies differ vastly in their quality and methodologies. A summary of these studies can be found in Appendix II. Some of these studies have looked at the ability to detect amblyopia and/or amblyopia risk factors in the general population (Mendez, 2015, Crescioni, 2015, Silverstein 2014, Ransberger, 2013) while others have focused on a clinical population that is more enriched in ocular disease (Silbert, 2014, Garry, 2014, Peterseim, 2014). The general population is defined as

population more similar to what would be found in a more "normal" screening environment and that is not as enriched in ocular pathology as that found in a pediatric optometric/ophthalmologic practice.

The majority of these studies have focused only on calculating sensitivity and specificity for the detection of amblyopia and/or amblyopia risk factors but have not included analysis on measurement agreement using Bland-Altman (Mendez, 2015, Crescioni, 2015, Silbert, 2014, Garry, 2014, Silverstein, 2014 Ransbarger, 2013) or ROC analysis. A few however, (Yajun 2016 and Crescioni 2015) have used ROC analysis. To date only one study Yajun et al has attempted to determine the optimal criteria setting based on ROC analysis. This study also investigated the measurement agreement of the Spot and GSE with Bland-Altman. One study Peterseim et al. (2014) examined the mean difference in measurement agreement between the Spot and GSE but did examine if this was consistent over the entire range of measurements.

The reason for referral in the Spot, based on various criteria settings, has been examined by multiple studies (Ransberger, 2013 Mendez, 2015, Silverstein). However, these studies did not determine if there was agreement in referral reason between the Spot and GSE.

Chapter 4: Methods and Objectives

4.1 Purpose, Aims and Objectives

The purpose of this study was to determine the accuracy of measurement outcomes of the Spot for hyperopia, astigmatism and anisometropia to those of GSE using Bland-Altman

and ROC analysis. Additionally, the sensitivity for the detection of amblyopia will also be examined. In order to improve the detection of the conditions discussed previously, the optimal referral criteria for the Spot will be determined. For this thesis the following aims and objectives were set:

1. Determine refractive error measurement agreement of the Spot with that of GSE for sphere, cylinder and anisometropia utilizing Bland-Altman analysis.

2. Determine sensitivity and specificity of the Spot for detection of amblyopia when utilizing the VIP pass/fail criteria, the pre-set Spot criteria v.1.1.50 and the modified Spot criteria.

3. Receiver operator characteristic analysis for hyperopia, astigmatism, anisometropia (both sph and cyl) and for all four factors combined.

4. Determine the optimal referral criteria of the Spot for detection of amblyogenic risk factors when specificity is set at 90%.

4.2 Methods

4.2.1 Subject Recruitment and Data Collection

Participants in the study were recruited from Head Starts and other Early Education Centers in the Boston area. Inclusion criteria comprised children aged 2 to 7 years and, for insurance purposes, who had not had an eye exam within the calendar year. Consents were

sent to school nurses to be handed out to all students. The protocol and informed consent were approved by the New England College of Optometry IRB and adhered to the tenets of the declaration of Helsinki. Only those subjects who had obtained proper parental consent were allowed to participate in the study. Subject data was collected from June 2012 through December 2013. Results from the Spot were downloaded and paired with GSE results using a unique patient identification code. A total of 483 subjects were screened and then examined by GSE.

4.2.2 Screening Protocol

Screening with the Spot was performed in an "as is" condition at each of the various schools and early education centers. Screening areas ranged from full size rooms to utility closet sized areas. Steps were taken at each location to reduce the ambient light levels as much as possible. In a previous pilot study on the testability of the Spot, a light meter was used to determine the ambient light levels at screening locations. In all cases light was reduced sufficiently to allow for vision screening. Screening with the Spot was performed by NECO students who were trained, by the manufacturers representative, in the use of the Spot. This training session was in-person and was performed prior to starting the study.

Subjects were brought to the screening location by the school nurse or other guardian and were assigned a unique identification code entered into the Spot along with general demographic data, including age, gender and state of optical correction. Any subject that was unable to be screened after three attempts was recorded as untestable and considered an automatic referral. After screening, subjects were sent for GSE on the mobile eye van.

4.2.3 Gold Standard Examination

Due to scheduling concerns, GSE was not always performed on the same day as vision screening but was always performed within a few weeks of the initial vision screening. All GSE exams were performed on the New England Eye Institute's mobile eye clinic by a pediatric optometrist. The mobile clinic consists of two separate exam lanes each 10 feet in length, fully equipped with electronic visual acuity charts, slit-lamp and binocular indirect ophthalmoscope (BIO). GSE consisted of visual acuity, entrance testing, simple binocular evaluation, cycloplegic retinoscopy and ocular health examination

Visual Acuity

Monocular and binocular visual acuity was measured using the Amblyopia Treatment Study (ATS –HOTV) protocol and digital chart at a distance of 10 ft. If the child was unable to verbally report the letters a matching card was used. The HOTV protocol was chosen instead of (ETDRS) logMAR chart because the ETDRS chart is often too difficult for children under the age of seven. Additionally, the ETDRS only has high test-retest reliability and good agreement with adults and children aged 7-12 years (Birch, 2009).

Entrance Testing and Binocular Evaluation

Entrance testing included extraocular motility, pupil testing, gross assessment of the anterior ocular structures, angles and adnexa and tactile pressures. Basic binocular evaluation included cover test (distance and near), near point of convergence, Hirschberg corneal reflex, Bruckner and Stereo vision testing with the Stereo Smile PASS Test 3. The Stereo Smile

PASS Test 3 is the most comprehensive set and includes the 480, 240, 120, 60 min of arc, demo and blank cards. Additionally, a 4 prism diopter BO test was performed on all subjects with unequal monocular acuities.

Assessment of Refractive Error

The subjects' refractive error was determined by cycloplegic retinoscopy using retinoscopy racks. Non-cycloplegic retinoscopy was also recorded but was not used for statistical analysis. Each subject was refracted after a minimum of 30 minutes after instillation of either two drops of combination drop 0.5% cyclopentolate, 2.5% phenylephrine and 0.25% tropicamide or two drops of Cyclogel (cyclopentolate hydrochloride 1%) in each eye. Glasses were then prescribed and ordered as part of the mobile van's outreach program if needed.

Anterior and Posterior Segment Evaluation

Anterior segment evaluation was grossly assessed and intraocular pressure was determined tentatively using the tactile method. If needed, a slit lamp and Goldman tonometry were available on the mobile clinic, but in no instances were these used. Posterior segment evaluation was performed in all subjects by BIO. Direct ophthalmoscopy was performed on those subjects who had poor cooperation for BIO.

4.3 Statistical Methods

4.3.1 Sensitivity and Specificity

Sensitivity and specificity are one approach to determine the diagnostic ability of a particular test. Sensitivity is the proportion of true positives that are correctly identified by the test while specificity is the proportion of true negatives that are correctly identified (Altman, 1994). Both sensitivity and specificity are independent of the population of interest subjected to the test. On the other hand, positive predictive value (PPV) and negative predictive value (NPV) which are also used to evaluate clinical tests are dependent on the prevalence of the disease in the population of interest. Determination of sensitivity and specificity is based on the following concepts:

1. True positive: the patient has the disease and the test is positive
2. False positive: the patient does not have the disease but the test is positive
3. True negative: the patient does not have the disease and the test is negative
4. False negative: the patient has the disease but the test is negative

Sensitivity is calculated as the number of TP/TP+FN. Thus a test that is 100% sensitive will correctly identify all patients with that particular disease. If a test is 70% sensitive will correctly identify 70% of those with the disease but will miss 30% of those with the disease. It would seem that having the highest sensitivity possible would be most desirable, however sensitivity needs to be balanced with specificity. Specificity is the ability of a test to correctly identify those patients without the disease and is calculated as

TN/TN+FP. So a test that is 100% specific correctly identifies all patients without the disease and a test that is 70% specific would correctly identify 70% of those without the disease but 30% of those without the disease would test positive (Lalkhen, 2008). For the purpose of this thesis, sensitivity and specificity were calculated utilizing Vasser Stats online clinical calculator 1: From an Observed Sample: Estimates of Population Prevalence, Sensitivity, Specificity, Predictive Values and Likelihood Ratios.

4.3.2 Receiver Operator Characteristic (ROC) Analysis

Although sensitivity and specificity are used routinely to help determine the accuracy of tests they are relatively crude measures and do not take into account cut-offs (pass/fail criteria) of a test. For example, as the cut-off point is raised the test becomes more specific but loses its sensitivity. Conversely, if the cut-off point is lowered then the test becomes more sensitive but not as specific. Receiver operator characteristic curves are used to take into account the cut-offs of a particular test and plot (1-specificity) against its sensitivity for all possible cut-off points. By taking the area under the curve (AUC) the overall accuracy of the test can be determined. An AUC value near 1.0 indicates both high sensitivity and specificity while a value near 0.50 would be the same as tossing a coin (Lalken, 2008).

4.3.3 Bland-Altman Analysis

Bland-Altman analysis allows for a relatively informal clinical interpretation of a new measurement method with that of a gold standard. It can be used to determine the bias (average discrepancy between methods) and the 95% limits of agreement (+/- 1.96 SD). If the bias is not close to zero then the two methods are producing systematically different

results. It is important to note that the bias is interpreted clinically, not statistically, as it may not be far enough from zero to cause problems clinically. The 95% limits of agreement indicate that for any future sample the difference between measures of the two methods should lie within the limits of agreement approximately 95% of the time. The width of the limits of agreement are also analyzed clinically not statistically. If the limits are narrow and the bias is small then the two methods can be used interchangeably. Trends in the data (proportional bias) can also be visualized by Bland-Altman analysis. For example, does the difference between the two measures increase as the average increases? (GraphPad, 2015). To evaluate the proportional bias formally, the difference between the two methods should be regressed on the average of two methods. (Bland, 1999, 1986).

Chapter 5: Results

5.1 Subject Characteristics and Data Overview

In total there were 483 subjects between the ages of 20 (1.67 y.o) and 99 (8.25 y.o) months who completed the study. There were 239 (49.5%) female and 244 (50.5%) male subjects. Approximately 50% of the study population was Hispanic. The other 50% of subjects were predominantly African American, followed by white and Asian American. Note that the demographics reported are only an estimation as the demographic section of the consent forms was not filled out by a majority of parents. The testability of the Spot was 100%.

Using the preset referral criteria v. 1.1.50, see table 2 for criteria, 108 (22.3%) subjects failed and 375 (77.6%) passed the screening. Note that the preset referral criteria uses spherical equivalent to determine pass/fail. The most common referral reasons were

astigmatism 18.63%, myopia 4.35%, anisometropia 2.28%, gaze 0.83% and anisocoria/hyperopia both at 0.41%. See figure 1. Please note that an individual subject could fail for more than one reason.

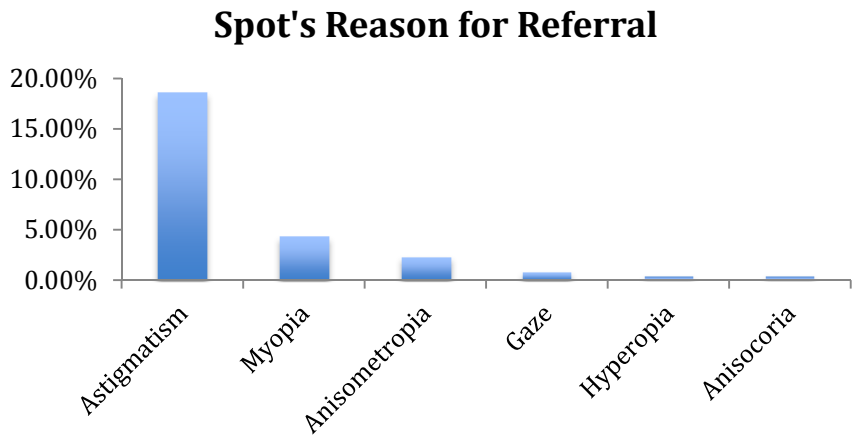


Figure 1: Reason for referral of Spot's pre-set criteria

Additionally, there were 27 (5.5%) confirmed cases of amblyopia. Criteria for the classification of amblyopia were based on the VIP criteria (VIP, 2004). Unilateral amblyopia was defined as a 2 line interocular visual acuity difference accompanied by strabismus and/or anisometropia. Those classified with bilateral amblyopia had reduced VA and/or an amblyogenic risk factor in each eye, astigmatism >-2.5D, hyperopia >5D, myopia > -8D. VA was considered reduced if it was below 20/50 in 3 year olds, 20/40 in 4 year olds and 20/32 in 5 and 6 year olds. Of those with amblyopia there were 30 diagnoses of refractive amblyopia, 9 (30%) anisometropic, 13 (43%) isoametropic and 8 (27%) meridional. Two

subjects had diagnosis of strabismic/microtropia amblyopia and there were no subjects with deprivation amblyopia. Five subjects had mixed cases of amblyopia resulting in more diagnoses than confirmed cases.

5.2 Measurement Agreement of Spot and GSE

To determine measurement agreement between the Spot and GSE, Bland-Altman analysis was performed. This was done by using the non-cyclopegic refractive error output of the Spot compared to the cycloplegic refractive error found in GSE. The refractive error, sphere and cylinder, was analyzed for the right and left eyes separately. Bland-Altman analysis was also performed for spherical anisometropia. All data for the analysis of Bland-Altman was performed using the raw output of the Spot without any formulation by the Spot's software. Spherical equivalence was not used to determine measurement agreement in any of the Bland-Altman analysis.

Bland-Altman analysis is interpreted clinically and not statistically. When examining the bias or average discrepancy between methods, the closer it is to zero the more likely that the two methods can be used interchangeably. The 95% limits of agreement indicate where the difference between measurement methods lie 95% of the time. A proportional bias (trend in data) can also be determined by regression and determines if the difference in measurement is the same over the entire range of measurements.

5.2.1 Bland-Altman Spherical Analysis

The bias (average discrepancy between methods) was 1.25 D for the right eye and 1.39 D for the left eye. The 95% limits of agreement for the right eye was upper 3.24 D and lower -0.74 D and left upper 1.40 D and lower -0.65 D. There was also a statistically significant, positive proportional bias (trend) in the data. This was determined by examining the difference between the two methods regressed on the average of the two methods. As the average of the two measures increases by one diopter, the bias increases by 0.51 D (P< 0.00) diopters for the right eyes and 0.57 (P<0.00) diopters for the left eyes. See figure 2 and 3.

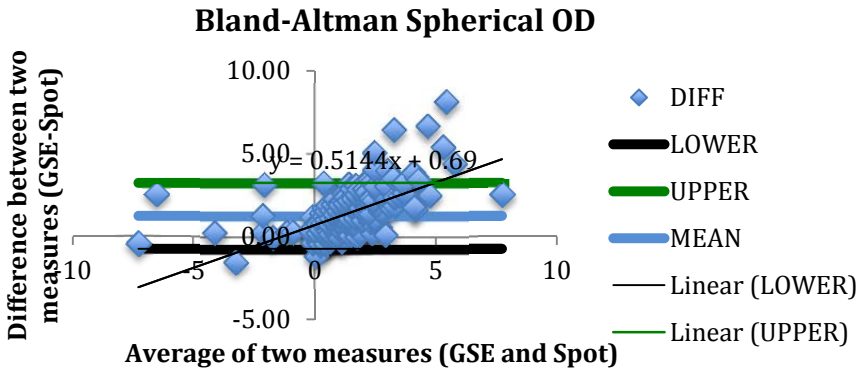


Figure 2: Bland-Altman Sph. OD; bias 1.25 OD, 95% LA OD upper 3.24 D, lower -0.74D, proportional bias OD 0.51 D (P<0.00),

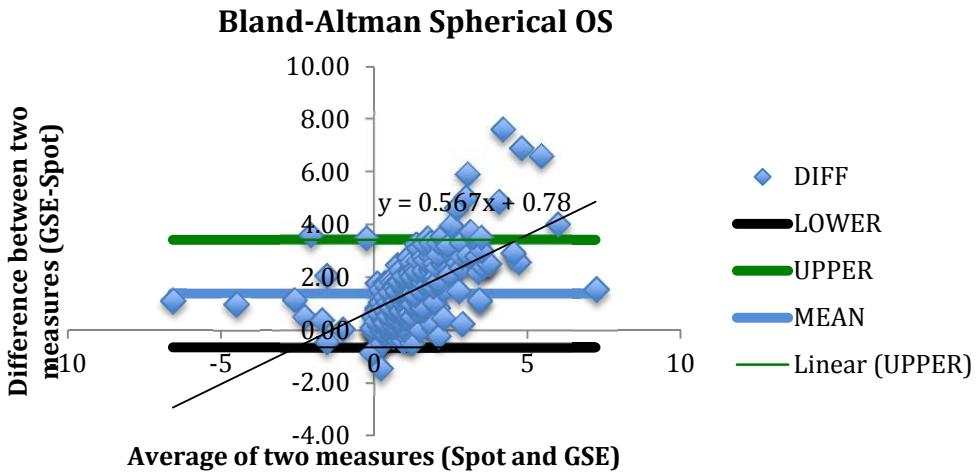


Figure 3: Bland-Altman Sph. OS, bias 1.39 D OS, 95% LA upper 1.40 D, lower -0.65 D, proportional bias OS 0.57 D (P<0.00)

5.2.2 Bland-Altman Cylindrical Analysis

The bias was 0.37 D for the right eye and 0.27 D for the left eye. The 95% limits of agreement for the right eye was upper 1.49 D and lower -0.75 D and left upper 1.40 D and lower -0.87. There was also a significant, proportional bias, trend, for the left but not the right eye. As the average of the two measures increases by one diopter, the bias increases by 0.03 (P<0.39) diopters for the right eyes and 0.13 (P< 0.00) diopters for the left eyes. See figure 4 and 5.

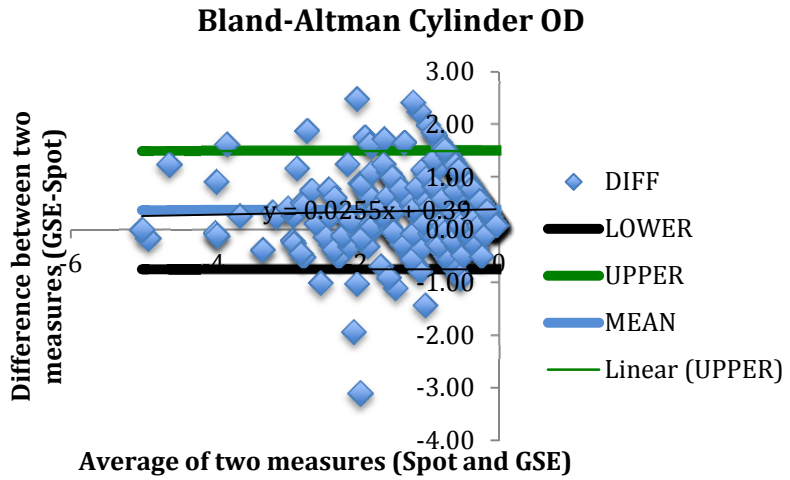


Figure 4: BA Cyl OD, bias 0.37, 95% LA upper 1.49, lower -0.75, proportional bias 0.03 (P<0.39)

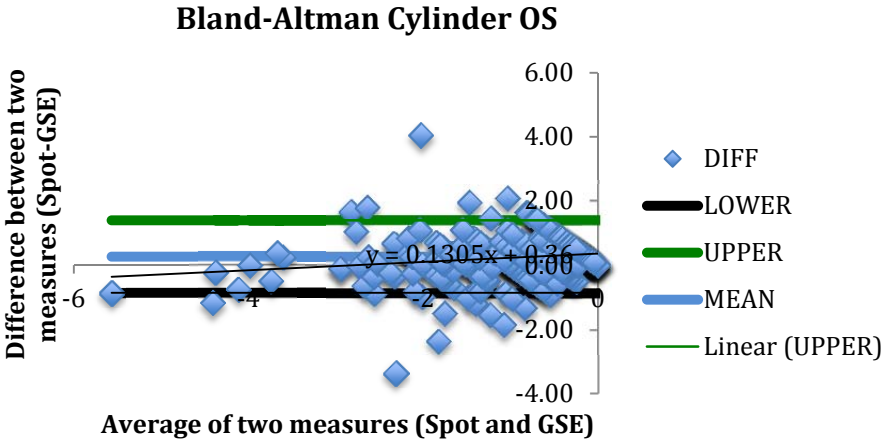


Figure 5: BA Cyl OS, bias 0.27, upper 1.40 D, lower -0.87 D, trend 0.13 (P<0.0)

5.2.3 Bland-Altman for Spherical Anisometropia

The bias was -0.03 D for spherical anisometropia. The 95% limits of agreement was upper 0.81 D and lower -0.87D. There was also a statistically significant, proportional bias (trend) in the data. As the average of the two measures increases by one diopter, the bias increases by 0.14 D (P< 0.00) diopters. See figure 6.

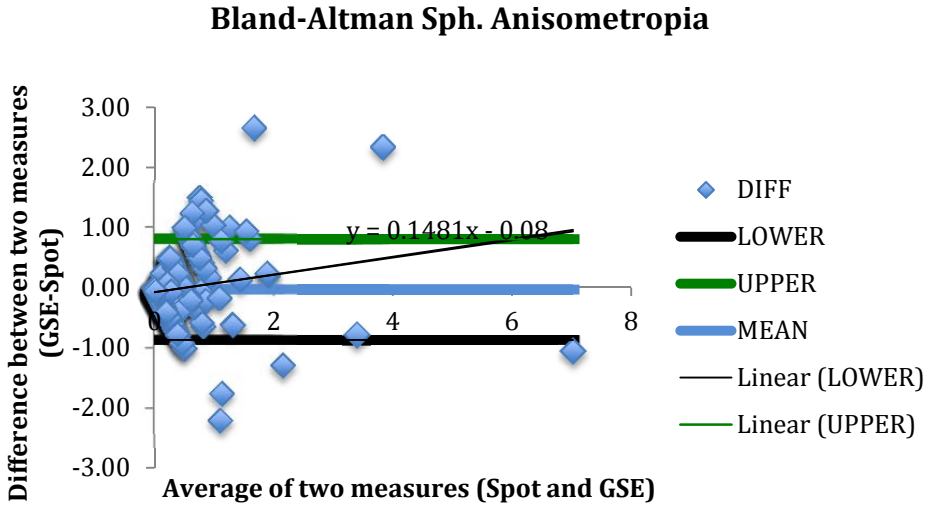


Figure 6: BA Sph. Aniso., bias -0.03 D, 95% LA upper 0.81 D, lower -0.87D, proportional bias 0.14 D (P<0.00).

5.3 High Hyperopic Subjects and Average Pupil Size Measurement

Out of 90 eyes with hyperopia >/= 3.5 D (based on cycloplegic examination) 41 eyes passed screening with the Spot, using the preset referral criteria v.1.1.50. Of those who failed screening with the Spot with this degree of hyperopia only two subjects had the correct referral diagnosis of hyperopia and two subjects had the referral diagnosis of myopia. To better understand the difference in measurements between the Spot and GSE in subjects

<>+3.50 D, regression analysis was applied to both groups. See figure 7 and 8. For both eyes the difference in measurement of the groups <+3.50 D and >/= +3.50 D were not statistically different from each other OD (P<0.15) and OS (P<0.53). However, for both groups <+3.50 D and >/= +3.50 D the difference in measurement was statistically different from zero. This was the same for both the right and left eyes. OD >/= +3.50 (P<0.00), <+3.50 (P<0.00) and OS >/= +3.50 (P<0.00), <+3.50 (P<0.00). The slope for the OD was 0.63 (P<0.00) for >/= +3.50 and 0.78 (P<0.00) for </= +3.50. The slope for the OS was 0.61 (P<0.00) for >/= +3.50 and 0.53 P<0.00 for < +3.50.

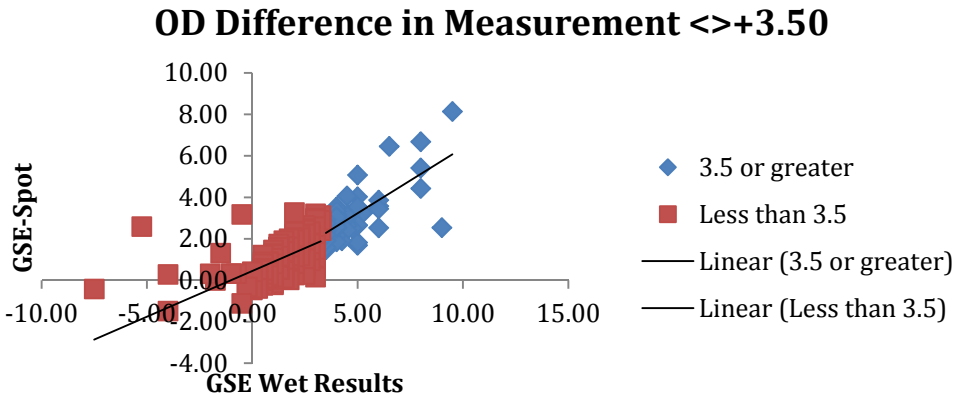


Figure 7: OD Slope 0.53 <+3.50 (P<0.00) , Slope 0.63 >/= +3.50 (P<0.00), <>+3.50 D diff. (P<0.15)

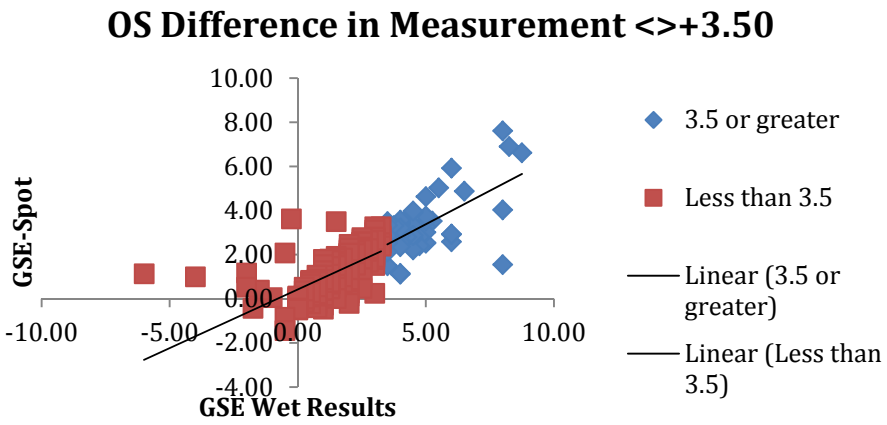


Figure 8: Slope 0.53 OS <+3.50 (P<0.00) , Slope 0.61 >/= +3.50 diff. (P<0.00), <>+3.50 D diff. (P<0.53)

In order to study the effects that accommodation might have on these two groups, the average pupil size was examined. For the OD <+3.50 the average pupil size was 6.50 mm and >/= +3.50 it was 6.28 mm and for the OS <+3.50 it was 6.50 mm and >/= +3.50 it was 6.33.

5.4 Sensitivity and Specificity for the Detection of Amblyopia

Sensitivity and specificity for the detection of amblyopia was calculated for three separate criterion, Spot's pre-set criteria v.1.1.50, VIP pass/fail criteria and a modified criteria developed from ROC analysis. The Spot uses spherical equivalence for determining a referral for the pre-set criteria and the pass/fail results of software v. 1.1.50 were used for analysis. For the VIP pass/fail criteria and modified criteria only the raw output of the Spot for sphere and cylinder was used. This was done to determine if the sensitivity and specificity would improve when using criteria that did not include spherical equivalence. It was

theorized that using spherical equivalent would underestimate hyperopia in subjects with high astigmatism and decrease the sensitivity for amblyopia detection. The agreement in reason for referral between the Spot and GSE was also examined for each of the three criteria. This was done to examine any mismatches in the reason for referral of the Spot and the cause of amblyopia according to GSE.

5.4.1 Spot Criteria Software v. 1.1.50

With the Spot's criteria v. 1.1.50 the Spot's sensitivity for detection of amblyopia was 77.78% CI (57.26%-90.62%) and the specificity was 80.44% CI (76.43%-83.92%). See table 4 for contingency table.

Table 4: Contingency table Spot Criteria v.1.1.50

	Amblyopia (GSE)	N0 Amblyopia (GSE)
Spot Fail	21 (TP)	89 (FP)
Spot Pass	6 (FN)	366 (TN)

The reason for referral according Spot's preset criteria in subjects that had amblyopia, was as follows; 2 failed for hyperopia, 7 for myopia, 2 for anisometropia and 18 for astigmatism. Note that a subject may be referred for more than one reason. For the subjects who passed screening with the Spot but who had amblyopia, 3 had hyperopia, 0 had myopia, 1 had strabismus, 1 had astigmatism and 3 had anisometropia. For comparison see table 5.

Table 5: Reason for referral for subjects with amblyopia preset

Reason for Amblyopia According to GSE					
Hyperopia	Myopia	Astigmatism	Anisometropia	Strabismus	Other
12	0	10	10	1	0
Reason for Referral According to Spot v.1.1.50					
2	5	17	2	0	1*
Amblyopia According to GSE but passed screening					
3	0	0	3	1	0

*anisocoria

5.4.2 VIP Criteria

When the VIP criteria for amblyogenic risk factors, see table 6, was applied to the raw refractive error output of the Spot the sensitivity for the detection of amblyopia was 77.77% CI (57.27, 90.62) and the specificity was 77.70% CI (73.53, 81.39). See table 7.

Table 6: VIP Amblyogenic Risk Factors

	Hyperopia	Myopia	Astig.
ISO.	>3.25D	>2 D	>1.50
Aniso.	>1D	>3D	>1.5D
Antimetropia	One eye >1D hyper and >1D between eyes	One eye >2D myopia and >3 D between eyes	

Table 7: Contingency table for VIP referral criteria

	Amblyopia (GSE)	No Amblyopia (GSE)
Spot Fail	21 (TP)	101 (FP)
Spot Pass	6 (FN)	352 (TN)

The reason for referral according to the VIP ARF criteria in those who had amblyopia was 3 for hyperopia, 0 for myopia, 0 for anisometropia and 19 for astigmatism. Note that a subject may be referred for more than one reason. For the subjects who passed screening with the VIP criteria but were determined to have amblyopia based upon GSE, 3 had hyperopia, 0 had myopia, 1 had strabismus, 1 had astigmatism and 3 had anisometropia. For comparison see table 8.

Table 8: Reason for referral in subjects with amblyopia (VIP Criteria)

Reason for Amblyopia According to GSE					
Hyperopia	Myopia	Astigmatism	Anisometropia	Strabismus	Other
12	0	10	10	1	0
Reason for Referral According to SPOT using VIP criteria					
3	0	19	0	NA	NA
Amblyopia According to GSE but passed Spot using VIP criteria					
3	0	1	3	1	0

5.4.3 Modified Criteria

When using the modified referral criteria determined by ROC analysis (hyperopia Sph 1.75 D, Cyl -1.50 D, anisometropia Sph 1.25 D and anisometropia Cyl 1.25 D) applied to the raw refractive error output of the Spot, the sensitivity of the Spot was 85.18% CI (65.39, 95.14) and the specificity was 77.04% CI (72.84, 80.78). See contingency table 9.

Table 9: Contingency table based on modified referral criteria

	Amblyopia (GSE)	No Amblyopia (GSE)
Spot Fail	23 (TP)	104 (FP)
Spot Pass	4 (FN)	349 (TN)

The reason for referral according to the modified criteria in amblyopic subjects was, hyperopia 10, myopia 1, anisometropia 7 and astigmatism 17. For the subjects who passed screening with the VIP criteria but were determined to have amblyopia based upon GSE, 2 had hyperopia, 0 had myopia, 0 had strabismus, 1 had astigmatism and 3 had anisometropia. For comparison see table 10. Note that a subject may be referred for more than one reason.

Table 10: Reason for referral for subjects with amblyopia (modified criteria)

Reason for Amblyopia According to GSE					
Hyperopia	Myopia	Astigmatism	Anisometropia	Strabismus	Other
12	0	10	10	1	0
Reason for Referral According to Spot using Modified Criteria					
10	1	17	7	NA	NA
Amblyopia According to GSE but passed Spot					
2	0	1	3	0	0

5.5 Receiver Operator Characteristic (ROC) analyses

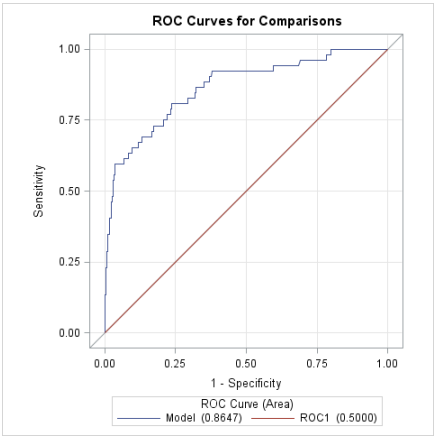
ROC analysis is used to determine the accuracy of a test using multiple criteria or cut-off points. This differs from sensitivity and specificity, which takes into account only one set of criteria. This is determined by plotting 1-specificity against the sensitivity for all possible cut-off points. By taking the area under the curve (AUC) the overall accuracy of the test can be determined. An AUC value near 1.0 indicates both high sensitivity and specificity while a value near 0.50 would be the same as tossing a coin. Analysis was performed using SAS statistical software for for both the right and left eyes separately and also for the eye with the highest refractive error (HRE). In the VIP study the eye with the highest refractive error for each subject was used for ROC analysis. For comparison, each eye was also analyzed separately.

To determine the optimal pass/fail criteria for the detection of amblyogenic risk factors specificity was set at 90%. Amblyogenic risk factors were defined as, hyperopia >/= +3.25, astigmatism >-1.50D, anisometropia sphere > 1D, cylinder > 1D. This procedure has been used previously in the VIP studies. There were only 13 subjects with myopia so these were omitted for analysis (VIP, 2004).

5.5.1 ROC for Right and Left Eyes

Using ROC analysis for each eye separately, the area under the curve (AUC) for hyperopia (defined as wet sph >/= +3.25 D) was 0.86 (0.80, 0.93) OD and 0.86 (0.81, 0.92) OS. See figure 9. Astigmatism (defined as wet exam >/= 1.50D) AUC was found to be 0.93 (0.90, 0.96) OD and 0.91 (0.87, 0.95) OS. See figure 10.

Hyperopia OD



Hyperopia OS

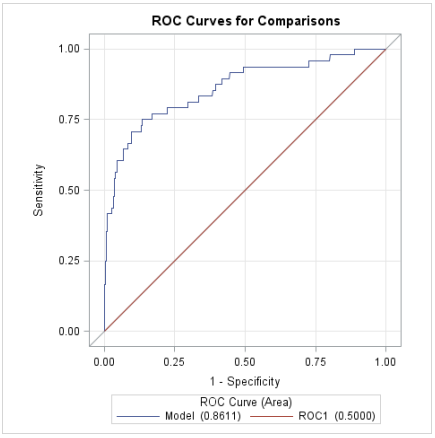
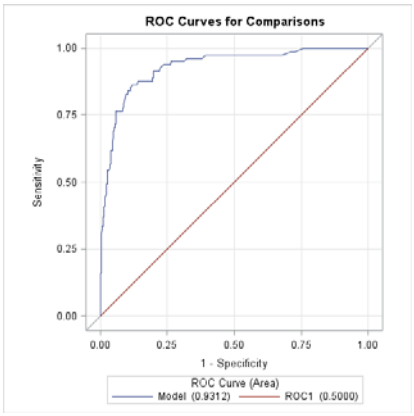


Figure 9: ROC hyperopia OD and OS: AUC OD 0.86 (0.80, 0.93) and OS 0.86 (0.81, 0.92)

Astigmatism OD



Astigmatism OS

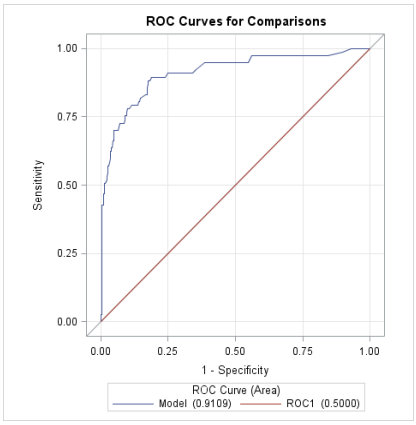


Figure 10: ROC for Astigmatism OD and OS: AUC OD 0.93 (0.90, 0.96) and OS 0.91 (0.87, 0.95).

5.5.2 ROC for Eye with Highest Refractive Error

When analyzing the eye with the highest refractive of each subject the AUC was as follows, most hyperopic meridian (0.74), cylinder power (0.85), anisometropia sph. (0.66), anisometropia cyl. (0.66) and with all four predictors combined (0.88). See figure 11.

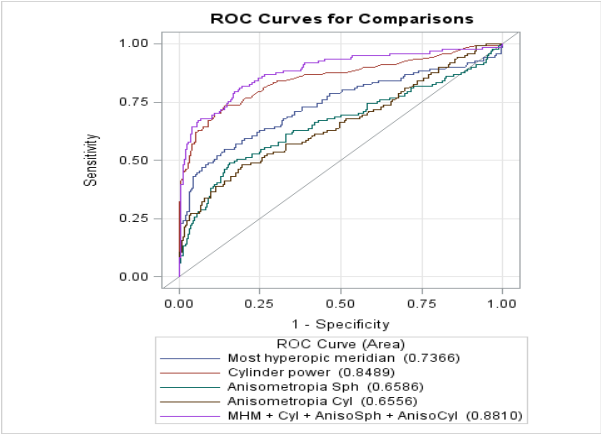


Figure 11: ROC curves for comparisons (HRE)

5.5.3 Modified Referral Criteria

When the specificity for the detection of amblyogenic risk factors was set at 90% the optimal pass/fail criteria of the Spot was determined. Amblyogenic risk factors were defined as hyperopia >/= +3.25, astigmatism >-1.50D, anisometropia sphere > 1D, cylinder > 1D. Due to the uniqueness of the study data, 21 unique criteria settings were found when the specificity was set at 90%. The multiple criteria settings can be found in appendix III. The highest sensitivity possible, 0.74% would result is a criteria setting of hyperopia Sph +1.00 D, Cyl -1.75 D, anisometropia Sph 1.25 D and anisometropia Cyl 1.50 D. The largest magnitude criteria that could be used while still maintaining a 90% specificity would be hyperopia Sph 3.00 D, Cyl -1.75, anisometropia Sph 1.75 and anisometropia Cyl 0.75. This however, would result in a sensitivity of 0.65%. The criteria setting that would allow for the largest magnitude criteria and that would maximize sensitivity is hyperopia Sph 1.75 D, Cyl -1.50 D, anisometropia Sph 1.25 D and anisometropia Cyl 1.25 D and would result in a sensitivity of 0.73%.

Chapter 6: Discussion

6.1 Data Overview

The testability of the Spot was 100% for this study and indicates that the device could be easy to use for large scale screening programs. This testability is similar to that found by Peterseim, (2014) testability of 100% and Crescioni (2015) who found a testability of 97%.

Using the preset referral criteria v.1.1.50 the most common reasons for screening

failure were astigmatism 18.63%, myopia 4.35%, anisometropia 2.28%, gaze 0.83% and anisocoria/hyperopia both at 0.41%. This is consistent with findings from Ransbarger (2013) who found the reasons for referral of the Spot to be astigmatism 2,393 (30.6%), ocular misalignment 879 (11.3%), anisometropia 90 (1.2%), myopia 82 (1.1%), hyperopia 63 (0.8%) and anisocoria in 16 (0.2%). The Ransbarger study used the 2003 JAAPOS screening guidelines that were programmed into the Spot.

Another study, done in Costa Rica, by Mendez (2015) examined the ability of the Spot, version 2.0.16, to detect amblyopia risk factors as defined by the 2013 AAPOS vision screening guidelines. The reason for referral based on the Spot's pre-set criteria in the Mendez study was 31 (14%) for astigmatism, 3 (1.4%) for anisometropia, 3 (1.4%) for hyperopia and 3 (1.4%) for myopia. Silverstein (2014) also examined the distribution of referral diagnosis by the Spot v.1.0.3 to be 51% for astigmatism, 8% for high refractive error and 14% for strabismus. See table 11 for a summary.

Table 11: Summary of Referral Reason for Spot

	Preset v.1.1.50 Current Study	Preprogramed 2003 AAPOS Criteria Ransbarger (2013)	Preset v.2.0.16 Mendez (2015)	Preset v.1.0.3 Silverstein (2014)
Astigmatism	18.63%	30.6%	14%	51%
Myopia	4.35%	1.1%	1.4%	NA
Gaze	0.83%	11.3%	NA	14%
Anisometropia	2.28%	1.2%	1.4%	NA
Anisocoria	0.41%	0.2%	NA	NA
Hyperopia	0.41%	0.8%	1.4%	NA
High Refractive Error	NA	NA	NA	8%

For all four studies the Spot referred at rates that were lower than expected for the population. In a study by the Multi-Ethnic Pediatric Eye Disease Study Group (MEPEDS), the rate of hyperopia was 26.9% in Hispanics, 25% for White Non-Hispanic, 20.8% for African American and 13% for Asian Americans (Wen, 2013). Additionally, a study in Iran found that the rate of hyperopia > +2.00 D was 20.5% for children entering school (Jamali, 2009).

Amblyopic Subjects

The population in this study had higher rates of amblyopia than expected, 5.5% vs. 1.6% to 3.6%, as reported in other studies (Simons, 2005). This may be due in part to the definition used to define amblyopia for this study, which was a modified version of the 2004 VIP definition of amblyopia. Unilateral amblyopia was defined as a two line interocular visual acuity difference accompanied by strabismus and/or anisometropia. Those classified with presumed bilateral amblyopia had reduced VA and/or an amblyogenic risk factor. These amblyogenic risk factors were defined as astigmatism >2.5D, hyperopia >5D, myopia > 8D. VA was considered reduced if it was below 20/50 in 3 year olds, 20/40 in 4 year olds and 20/32 in 5 and 6 year olds.

The majority of presumed amblyopic subjects had refractive amblyopia and two subjects also had a diagnosis of strabismus/microtropia. Interestingly, the most highly cited diagnosis for refractive amblyopia was isoametropic 43%, followed by anisometropic 30% and meridional 27%. This is in contrast to other studies that indicate that isoametropic amblyopia is rare and accounts for only 1-2% of refractive amblyopia cases and that anisometropia and/or strabismus accounts for 90% of amblyopia cases (Arnos, 1987, AOA, 2004). This is likely due to the definition of amblyopia used in this study and that the diagnosis of amblyopia was based on only one visit. Due to scheduling difficulties the mobile eye clinic is unable to offer complete follow-up services. This means that the impact on the visual acuity from first time glasses wear could not be established. Another possibility is that this study population was more enriched for amblyopia than the general population. It is

likely that this is due to school/parent selection of those subjects who received screenings and exams. Although most schools tried to have all children participate, perhaps those parents who signed consents were those who had noticed potential visual problems or had a family history of vision problems and were thus more inclined to send their child for an exam.

6.2 Spherical Measurement Agreement

Bland-Altman analysis, compares a new measurement method with that of a gold standard. For right and left eyes spherical component, the bias (average difference between measures) was 1.25 D for the right eye and 1.39 D for the left eye. This indicates that the GSE averages 1.25 D more plus for the right eye and 1.39 D more plus for the left eye. In other words the Spot underpluses sphere. Bias is interpreted clinically and not statistically and it is debatable if this is a clinically significant difference. The closer to zero the bias is, the more likely it is that the two measures can be used interchangeably. This is consistent with other studies (Peterseim, 2014) that have investigated the mean difference in sphere measurements for the Spot and cycloplegic exam and found the Spot to over minus by -1.35D. However, just examining the mean difference in measurement between the Spot and GSE does not determine if there is equal agreement throughout the range of measurements. This can only be done by further analysis by Bland-Altman. Yajun (2016) also found that the Spot had a tendency to underestimate hyperopia and to overestimate myopia by about 0.49 D when using Bland-Altman analysis. It should be noted that the Yajun study did use spherical equivalence to determine this shift.

The 95% limits of agreement for the right eye was upper 3.24 D and lower -0.74 D and for the left eye it was upper 3.43 D and lower -0.65 D. This represents an arguably wide range clinically. Two methods of measurement can be considered essentially equivalent if the limit range is narrow. This finding is similar to Yajun (2016) who also found wide limits of agreement for the Spot with Bland-Altman.

There was also a statistically-significant proportional bias (trend) for both the right and left eyes of 0.51 D P< 0.00 and 0.57 D P <0.00 respectively. This indicates that the methods do not agree equally throughout the range of measurements. As the magnitude of hyperopic refractive error increases, the measurement agreement of the Spot decreases. Thus, the GSE and the Spot could be used interchangeably for low- level refractive error, ~ < 3.50 D only. As the Spot is used as a screening device to target higher levels of refractive error, the clinical relevance of a high measurement agreement for low-level hyperopic refractive error, is debatable. This finding is also consistent with the findings by (Yajun, 2016), who found an increasing difference between measures at ~ +3.50 D with Bland-Altman.

6.3 Cylindrical Measurement Agreement

Bland-Altman analysis, see figures 4 and 5, was also applied to the cylindrical component of both right and left eyes. The bias (average difference between measures) for the right eye was 0.37 D and 0.27 D for the left. This represents great measurement agreement as the closer to zero the bias is the more likely it is that the two measurement agreements can be used interchangeably. This bias indicates that the GSE averages 0.37D more plus in cyl for the right eye and 0.27 D more plus for the left than the Spot. This differs

slightly from the results found by Peterseim (2014) who found that the mean difference between measures of the Spot and GSE for cylinder was 0.36 D. This means that the Spot overplused cylinder by 0.36 D compared to cycloplegic exam. Additionally, Yajun (2016) also found moderate agreement for cylinder measurement with Bland-Altman analysis. As discussed previously, the bias does not determine if there is equal agreement throughout the range of measurements, which is done by further analysis by Bland-Altman.

The 95% limits of agreement OD (1.49 D, -0.75 D) and OS (1.40 D, -0.87 D) could be considered to be clinically large. This is also consistent to Yajun (2016) who found the 95% limits of agreement to be large. In terms of a proportional bias (trend), there was a statistically significant trend for the left eye only. The proportional bias was OD 0.03 D ($P < 0.39$) and OS 0.13 D ($P < 0.00$). However, this does not represent a clinically significant difference. There would have to be a very large degree of astigmatism for a clinically significant difference in measurement agreement ie. 10D of astigmatism would have a 1.3D difference in measurement agreement. This is also consistent to the findings by (Yajun, 2016) who found that there was better agreement for the J0 and J45 meridians between the Spot and GSE.

6.4 Spherical Anisometropia Measurement Agreement

Bland-Altman analysis of spherical anisometropia showed a bias (average difference between measures) of -0.03 between eyes. See figure 6. This represents good agreement between the two measurement methods. The 95% limits of agreement were upper -0.87 D and lower 0.81 D and can be considered an acceptable range clinically. There was a

statistically significant proportional bias (trend) of 0.14D P<0.00. However, this is not really clinically significant, as the magnitude of anisometropia would have to be very great in order for the bias to have effect. To date there is no equivalent analysis, by Bland-Altman, of anisometropia in the literature. This analysis is needed to determine how close the measurement agreement of the Spot and GSE are and additionally to determine if there is any trend in the data over the range of measurement errors.

6.5 Sensitivity and Specificity for Detection of Amblyopia

Pre-Set Criteria Software Version 1.1.50

Using the pre-set criteria v.1.1.50 for the detection of amblyopia, the sensitivity of the Spot was 76.92% CI (55.91, 90.25) and the specificity was 80.26% (CI 76.25, 83.76). This differs to a study by (Garry, 2014) who found the sensitivity for the detection of amblyopia, Spot v.1.0.3, to be 92% and 41% specific when using the 2013 AAPOS criteria and the sensitivity to be 89% and specificity 53% for Spot v.1.1.51. The Garry (2014) study also examined the sensitivity and specificity of the Spot for the detection of amblyopia risk factors (ARF) and found it to be 89% sensitive and 71% specific for Spot v.1.0.3 and 85% sensitive and 88% specific for Spot v.1.1.51 (Garry, 2014). This is comparable to another study by (Silbert, 2014) that reported a sensitivity of 80% and specificity of 74% for the detection of ARF. Neither of these studies however used ROC analysis to determine the sensitivity and specificity of the Spot over a range of cutoffs.

Although the sensitivity and specificity for the detection of amblyopia of the Spot appeared to be good, there was little agreement on the referral reason between the Spot and

GSE. See table 5. This is particularly true for those individuals with high hyperopia. There is much better agreement for those with astigmatism. In conclusion, this indicates that while the Spot has a relativity high overall sensitivity and specificity for the detection of amblyopia the reason for referral does not always correspond to what is found on GSE. With the current pre-set criteria v.1.1.50, amblyopic individuals with high hyperopia but low degrees of astigmatism may be missed by the Spot.

Sensitivity and Specificity utilizing the VIP Criteria

Since the Spot has the ability for a screener to change the referral criteria programmed into the device, the sensitivity and specificity for the detection of amblyopia based on the VIP criteria, was analyzed. However, it should be noted that the Spot currently uses spherical equivalent to refer for the spherical component of the refractive error. In order to truly change the referral criteria on the Spot to the VIP criteria this would need to be changed by the manufacturer. For this reason the VIP criteria was applied to the raw refractive error output of the Spot in order to determine sensitivity and specificity.

When applying the VIP criteria, see table 6, to the raw refractive error output of the Spot. The sensitivity of the Spot was 77.77% CI (57.27%, 90.62%) and the specificity was 77.70% CI (73.53%, 81.39%) for the detection of amblyopia. This is comparable to a study by (Silbert, 2014) that reported a sensitivity of 74%. As with the pre-set criteria v.1.1.50 there was little agreement in terms of referral reason for hyperopia but good agreement for astigmatism. See table 8. There was also little agreement for anisometropia. However, the referral agreement for myopia was better than the preset criteria v.1.1.50. This was most

likely due to not using spherical equivalence, which could overestimate myopia in high hyperopes. In conclusion, the results of using the VIP criteria for referral were similar to those using the preset criteria v.1.1.50 with the notable exception of myopia.

Sensitivity and Specificity utilizing the Modified Referral Criteria

An optimal referral criteria to detect amblyogenic risk factors when specificity is set at 90% was determined by ROC analysis. This criteria was as follows, hyperopia Sph 1.75 D, Cyl -1.50 D, anisometropia Sph 1.25 D and anisometropia Cyl 1.25. Again due to the fact that the Spot uses spherical equivalent as part of its software to determine a referral this criteria was applied to the raw refractive error output of the Spot in order to determine sensitivity and specificity for the detection of amblyopia. In terms of referral reason, there was relatively good agreement in referral reason for hyperopia, myopia, astigmatism and anisometropia. See table 10. Using the modified criteria increases the sensitivity of the Spot for the detection of amblyopia while also increasing the agreement of referral reason between the Spot and GSE. Additionally, the specificity for the detection of amblyopia remains similar to the other referral criteria examined.

Another study by (Yajun, 2016) also developed a modified referral criteria based on ROC analysis to detect amblyopia risk factors. They determined that the setting the optimized referral criteria for hyperopia to +2.375 D the sensitivity for the detection of amblyopic risk factors increased from 79.66% to 98.31% with minimal sacrifice of specificity. For astigmatism and anisometropia, the optimized criteria were 1.25D and 1.125D respectively. For myopia it was determined to be -2.0D. Both the modified criteria

proposed by this study and that of (Yajun, 2016) are lower than that currently preset on the Spot. Additionally, this modified referral criteria is also lower than what is recommended by both the VIP study and AAPOS. Thus it is not recommended to simply change the referral criteria of Spot based on professional referral recommendations because of the Spot's limitations in measurement agreement.

6.6 ROC Analysis for Right and Left Eyes

Receiver operator characteristic curve analysis allows for a more detailed analysis of the accuracy of a test than sensitivity and specificity as it takes into account all possible cut-offs (criteria) and the resulting sensitivities and specificities associated with those cut-offs. By taking the area under the curve (AUC), the overall accuracy of the test can be determined. An AUC value near 1.0 indicates both high sensitivity and specificity while a value near 0.50 would be the same as tossing a coin.

For both the right and left eyes for the detection of hyperopia the AUC was 0.86 (0.80, 0.93) and 0.86 (0.81, 0.95). This is lower than the AUC reported by (Yajun, 2016) who reported an AUC for the detection of hyperopia of .995. The Yajun study did not explain which eye or if both eyes were chosen for analysis. Additionally, the AUC for hyperopia may also differ from this study due to differences in the study populations. The Yujun study was undertaken in a population of school aged Chinese children. In terms of comparison to other devices on the market, (Kulp, 2014) found the AUC when specificity is set at 90% of the Suresight vision screener for hyperopia detection to be 0.85, Retinomax autorefractor 0.91

and non-cylcoplegic retinoscopy 0.92. See table 12. In conclusion, it appears that the Spot is similar to the SureSight in hyperopia detection for each eye individually.

For the detection of astigmatism the AUC for the right and left eyes was 0.93 (0.90, 0.96) and 0.91 (0.87, 0.95). This is similar to the AUC reported by (Yajun, 2016) who reported an AUC of 0.90 for the detection of astigmatism and (Crescioni, 2015) who found an AUC of 0.97. In terms of comparison to other devices on the market (Kulp, 2014) found the AUC of the SureSight for astigmatism detection was 0.95, Retinomax autorefractor AUC was 0.95 and non-cycloplegic retinoscopy was 0.90. Thus the Spot performs similarly to non-cycloplegic retinoscopy for the detection of astigmatism.

6.7 ROC Analysis for Eye with Highest Refractive Error (HRE)

Results from the ROC analysis for the Spot utilizing the data from the eye with the highest refractive error shows that for hyperopia and astigmatism the AUC was 0.74 (0.68, 0.80) and 0.85 (0.80, 0.90) respectively. The AUC for both spherical and cylindrical anisometropia was 0.66. This is much lower than the AUC reported by Yajun (2016) who reported an AUC of 0.995 for hyperopia, 0.90 for astigmatism and 0.92 for anisometropia. This is likely due to differences in how the data was analyzed. For this study, the eye with the highest refractive error out of each subject pair was chosen for analysis, which is similar to the protocol used in the VIP studies. The differences in AUC may also be due to variations in the study populations.

The AUC for the eye with the highest refractive error is also lower than when the eyes are analyzed separately for both hyperopia and astigmatism. This may result from the

difficulty of the Spot to detect higher refractive errors, which are selected for when analyzing binocularly the eye with the highest refractive error. Compared with other devices on the market and with NCR, the Spot underperforms in all four categories; most hyperopic meridian, cyl power and anisometropia (Kulp, 2014). See table 12 for comparison.

Table 12: Comparison of device AUC

Refractive Error Combined	**AUC**
Spot	0.88
NCR (ref)	0.92
Retinomax Layscreener (Ref)	0.92
Suresight Layscreener (Ref)	0.92
Hyperopia	
Spot HRE*	0.74 (0.68, 0.80)
Spot Right Eyes	0.86 (0.80, 0.93)
Spot Left Eyes	0.86 (0.81, 0.92)
NCR	0.92
Retinomax Layscreener	0.90
Suresight Layscreener	0.87
Yajun et. al Spot	0.995
Astigmatism	
Spot HRE	0.85 (0.80, 0.90)
NCR	0.90
Spot Right Eyes	0.93 (0.90, 0.96)
Spot Left Eyes	0.91 (0.87, 0.95)
Retinomax Layscreener	0.95
Suresight Layscreener	0.95
Yajun et al. Spot	0.90
Anisometropia*	
Spot Sph HRE	0.66
Spot Cyl HRE	0.66
NCR	0.82
Retinomax Layscreener	0.78
Suresight Layscreener	0.90
Yajun et al Spot	0.92

*note anisometropia (cyl and Sph is combined for NCR, Retinomax and Suresight
*HRE (highest refractive error eye)

6.8 Recommended Optimal Criteria Guidelines for Spot

Using ROC analysis to determine optimal referral criteria when specificity is set at 90% has been used previously in the VIP studies (VIP, 2014). Using a specificity set at 90% allows for a balance of sensitivity and specificity to reduce over referrals. When using this method there were 21 outcomes for referral criteria due to the uniqueness of the study population. See Appendix III. Upon examination of these criteria options, the criteria setting that would maximize sensitivity and that would allow for the largest magnitude criteria to be used was selected. These criteria are as follows; hyperopia Sph 1.75 D, Cyl -1.50 D, anisometropia Sph 1.25 D and anisometropia Cyl 1.25 D. The modified critera would result in a sensitivity of 73% for the detection of amblyogenic risk factors. Another criteria setting has been proposed by (Yajun, 2016) discussed previously. Both the modified criteria proposed by this study and that of (Yajun,2016) are lower than that currently set on the Spot and that is recommended by both the VIP study and AAPOS.

6.9 Pupil Size and Accommodation

To determine if lack of accommodation control impacted the ability of the Spot to accurately measure hyperopia, the average pupil size of two subject groups was analyzed. These groups were individuals with <+3.50 D hyperopia and >/= +3.50 D of hyperopia as determined by cycloplegic retinoscopy. It was theorized that the average pupil size would be significantly smaller in individuals with >/= +3.50 D of hyperopia due to accommodation. In subjects with < +3.50 D of hyperopia the average pupil size was OD 6.50 mm and OS 6.50mm. . In subjects with >/= +3.50 D the average pupil size was OD 6.28 mm and OS 6.33

mm. The Spot reports an accuracy in pupil measurement of +/- 0.1 mm and a range of 4.0 mm to 9.0 mm in 0.1 mm increments. However, further work needs to be done to determine the accuracy of the Spot in measuring pupil size. Thus it seems that there is no clinically significant difference in the average pupil size between the two groups. This suggests that accommodation may be accurately controlled even in subjects with high degrees of hyperopia. However, pupil size in children may also not be a reliable method of determining if accommodation is happening.

The limitations of the Spot to adequately measure hyperopia may be due to inherent problems in photoscreening technology which has been noted on previous studies. Frank Schaffel in Germany developed the initial photoscreening device, the PowerRefractor, for use in animal model research. This developed into the PowerRefractor II, which became a commercially available photoscreening device which was later bought and developed into the PlusoptiX and later the PlusoptiX II. Photoscreening technology similar to that used in the PlusoptiX II was also used in the Spot device.

In a study by Satiani and Mutti, screening for high hyperopia in infants using the PowerRefractor, it was noted that there was a significant underestimation of higher-level hyperopia by the PowerRefractor when compared to cycloplegic retinoscopy. This occurred despite the subjects being under cycloplegia for both screening and retinoscopy. They found that when +4.50D accessory glasses were worn in subjects who had initial readings >/= +2D the underestimation of hyperopia improved. The AUC of ROC was examined and found to be 0.69 to 0.50 when no glasses were worn and improved to 0.87 to 0.98 when the glasses were worn with cut-points between +3.50 and +5.00 with retinoscopy (Satiani, 2011).

Another study on the earlier photoscreening technology of the PowerRefractor by Schimitzek (2002), also noted that without cycloplegia the PowerRefractor, tended to underplus refractive error when compared to cycloplegic retinoscopy. Similarly to the results of the current study, the PowerRefractor yielded good results from the spherical component of low-level refractive error, but ametropia was underestimated outside +/- 4 D. (Schimitzek, 2005).

In another study by the same author examining the efficiency of the PowerRefractor as a screening device for amblyogenic factors found that the device missed several cases of anisometropia and high magnitude astigmatic and hyperopic ametropia. It was also noted that outside a range of -7.0 D to +5.0 D the device could even yield a result of emmetropia. (Schimitzek, 2002). This is a similar result to the results of this study where cases of high hyperopia were missed and assigned lower readings rather than being considered out of range.

Another study by Mirzajani et al. compared the results of the Plusoptix S08, another photoscreening device developed after the PowerRefractor and with technology similar to the Spot. The study used Bland-Altman analysis to conclude that higher amounts of hyperopia were underestimated by the PlusOptix S08 photoscreener compared to cycloplegic examination (Mirzajani, 2013).

6.10 Study Limitations

Limitations for this study included a relatively small sample size, which limited the ability to analyze all types of anisometropia and significant refractive error as categorized by

VIP. Additionally, due to the small sample size there were relatively few amblyopes (27 in total), which may have affected the results of the sensitivity and specificity for the detection of amblyopia. This study also did not examine the accuracy of the gaze (strabismus) detection capabilities of the Spot or the difference in pupillary size (anisocoria) that are included as part of the pass/fail criteria of the Spot.

6.11 Summary and Further Work

In conclusion, the Spot appears accurate in detection of lower level refractive errors and astigmatism. However, the Spot has difficulty in the detection of higher magnitude hyperopia as is evidenced by the results of Bland-Altman analysis. According to ROC analysis the Spot also performs relatively poorly for hyperopia compared to other devices on the market.

For the detection of amblyopia as determined by gold standard examination, the sensitivity and specificity of the Spot seems to be quite good. However, upon examination of the reason for referral there is a mismatch between the reason for referral of the Spot and the reason for amblyopia as determined by GSE for the pre-set criteria v.1.1.50 of the Spot. This mismatch improves somewhat when the spherical equivalent is not used for determining referral. This is supported by the increase in myopia referral agreement when the VIP criteria were applied to the raw refractive error output of the Spot. The mismatch in referral reason improved further when the optimal referral criteria were applied to the Spot's raw refractive error output.

In order to increase the accuracy of the Spot device, an optimal referral criteria was determined by ROC analysis. This optimal referral criteria increased both the Spot's

sensitivity for the detection of amblyopia and increased the agreement in referral reason of the Spot and GSE. Thus it is recommended that the referral criteria of the Spot be adjusted to reflect the modified criteria and that spherical equivalent should no longer be used for the referral. Further work is needed to examine the accuracy of the Spot for the detection of gaze misalignment, pupil size and comparison to non-cycloplegic retinoscopy.

Appendix I: Summary of the American Association for Pediatric Ophthalmology and Strabismus AAPOS State-by-State Vision Screening Requirements as of May 2, 2011

State	Mandated Vision Screening	Testing Used	Ages Screened
Alabama	General Medical Public School	N/A	N/A
Alaska	Yes	Specified by District	Specified by District
Arkansas	Yes	N/A	Pre-K, 2, 4,6, 8 and transfer students
Arizona	No State Policy	NA	Recommended annually to 6th grade
California	Yes	NA	Upon Enrollment and every 3rd year until completion of grade 8
Colorado	Yes	NA	K-3, 5, 7, 9
Connecticut	Yes	NA	K-6 annually and 9th
Delaware	Yes	NA	K, 2, 4, 7, 9, 10 by Jan. 15th, new enrollees, special education students, students referred by teacher or administrator
DC	Yes	NA	Within 150 days prior to entering Pre-K, K, 1,3,5,7,9,11
Florida	Yes	NA	K, 1,3,6 and new enrollees
Georgia	Yes, public school	NA	1st grade
Hawaii	No	NA	NA
Idaho	No	NA	NA
Illinois	Eye Exam and screening (public, private or parochial)	NA	EE: prior to K or entering school the first time VS: all preschool, K, 2, 8, Special Education, referred by teachers and transfer students Recommended 4, 6, 10 and 12
Indiana	Yes	Acuity	Annually 3, 8 And before enrolling K or grade 1
Iowa	No	NA	Recommend parents send child for exam by age 7. Appropriate assessment prior to enrollment in special education
Kansas	Yes, public and nonpublic	NA	During 1st year of admission and not less than once every two years
Kentucky	Preventive health care examination including vision (public or head start)	NA	Evidence of vision examination must be submitted no later than Jan 1 of the 1st year a 3, 4,5 or 6 year old is enrolled
Louisiana	Yes	Test sight, color screening	Entering K and 1st grade within 30 days of admission
Maine	Yes	Specification provided	K,1,3,5,7,9 all new and transfer students
Maryland	Yes (religious	Room shall be provided	First year of entry, 1st, 8th or 9th

	exemption available)	with specific lighting for test used	
Massachusetts	Yes (Public)	NA	All public school students annually
Michigan	Yes	NA	Requires proof that vision has been tested since age 3 prior to entering kindergarten
Minnesota	Yes, public	NA	At least once before school entrance (K or 1st)
Missouri	CEE	Protocols listed	By optometrist of physician by Jan 1 of the first year child is enrolled in school
Mississippi	Yes	School Nurse	NA
Montana	No	NA	Recommended evaluation on periodic basis
Nebraska	Yes	Evaluation by physician, physician assistant, advanced practical nurse or optometrist	6 mos prior to beginning the 2006-2007 school year and every year thereafter
New Hampshire	No	No	Recommends annually
North Carolina	Yes	NA	Within 180 days of the start of the school year, for those that fail a CEE is required funds are available for parents
North Dakota	No	NA	NA
Nevada	Yes	NA	Two years elementary, one grade middle school, one grade at high school
New Mexico	Yes (parents can opt out)	NA	Pre-K, K, 1st and 3rd, any new or transfer
New Jersey	Yes	Visual Acuity	Biennial screening K through grade 10
New York	No, subject to regulations of the commissioner	NA	NA
Ohio	Yes	Methods must be approved by the department of health	Prior to Nov. 1 of year enrolled in K or 1st grade
Oklahoma	Yes		Proof of screening during previous 12 mos. For public K, 1st, 3rd, LEP must send written report
Oregon	Yes	NA	NA
Pennsylvania	Yes	By nurse, teacher or medical technician, Snellen chart	NA
Puerto Rico	NA	NA	NA
Rhode Island	Yes	NA	K, 1,2, 3, 4,5,7 and 9th
South Carolina	No	NA	Recommended K, 1st, 3rd, 5th, 7th, 9th, 11th
South Dakota	No	NA	NA
Tennessee	Yes	NA	Yearly Pre-K, K, 2nd, 4th, 6th and 8th
Texas	Yes (public and private)	NA	Age 4, K entrants, 1st, 3rd, 5th and 7th
Utah	Yes	NA	Under age 8, screenings are free
Vermont	Yes	PCPs and school districts based upon research guidelines	1st, 2nd, 3rd, 5th, 7th and 9th
Virginia	Yes	Principal must ensure testing	NA
Washington	Yes	Must use Snellen for	K, 1, 2, 3, 5, 7 and if resources permit all other grades

		distance acuity (properly illuminated and glare free)	
West Virginia	Yes	NA	All pre-K must have healthcheck on file which includes vision, all new school enterers must be screened prior to enrollment
Wisconsin	Eye Exam	NA	Each pupil entering K provide evidence that the eyes have been examined by an optometrist or licensed physician. Providers are encouraged to provide free exams to those in need
Wyoming	Yes	Qualified personnel	NA

***based upon data complied by the American Association for Pediatric Ophthalmology and Strabismus. Accessed June 2, 2015. http://www.aapos.org/resources/state_by_state_vision_screening_requirements/**

Appendix II: Summary of Spot Studies

Authors	Purpose	N Size	Methods	Results	Limitations
Silbert, David MD, FAAP and Matta, Noelle CO, CRC. *Journal of AAPOS 2014*	Compare Spot's detection of amblyopia risk factors to 2013 AAPOS referral criteria	151 children ages 1-6 y.o from an Ophthalmology clinic with high rates of pathology	Spot referrals were compared to results from CEE with cycloplegia using APPOS 2013 referal criteria	With AAPOS '13 criteria 65% were found to have amblyopia risk and Spot agreed with this referral criteria 85% of the time.	Small n size, a retrospective study design, cycloplegic refractions performed up to 6 mos prior to vision screening and utilizing sensitivity, specificity, PPV and NPV which are a relativity crude assessment of accuracy
Garry, Glynnis BA, Donahue, Sean MD, PhD. *Journal of AAPOS 2014*	Determine the sensitivity and specificity for two different pre-set referral criteria (v1.0.3) and (v1.1.51) to detect amblyopia risk factors and amblyopia	155 ages 2-9 y.o from an Ophthalmology clinic with high rates of pathology	Subjects were screened with the Spot prior to receiving a CEE with cycloplegia	For detection of amblyopia risk factors v1.0.3 was 89% sensitive and 71% specific and v1.1.51 was 85% sensitive and 88% specific. For detection of amblyopia v.1.1.51 was 92% sensitive and 41% specific and v1.1.51 was 89% sensitive and 53% specific.	Small n size, a cohort that was rich in ocular pathology and that might not represent the general screening population and again using the relativity crude measures of sensitivity and specificity. Additionally, the type of amblyopia, refractive, strabismic or occlusive was not reported and directly compared to the results from the Spot.
Ransbarger, Kristiane MD, Dunbar, Jennifer A. MD, Choi, So Eun BA, Khazaeni, Leila M. MD. *Journal of AAPOS 2013. (REF #)*	Describe results of community vision screening with the Spot photoscreener stratified by age group	8,317 subjects with matching CEE in 300 subjects ages to 72 mos.	Children who failed screening were referred for a CEE with cycloplegia from various providers and results of screening were compared to GSE utilizing the same referral criteria AAPOS 2013 stratified by age group	Reason for referral was confirmed in 55.7% of the CEE. PPV was for any amblyogenic risk factor was 65.7%. PPV for astigmatism was 73.5% and for ocular misalignment 52.0% across all age groups. PPV could not be calculated due to small n size for anisometropia, myopia, hyperopia and anisocoria. There was no record for 83.6% of children who were referred from photoscreening.	Only 300 examination records for 2,393 children screened and referred by the Spot, no GSE on children who passed Spot, no examination of Spot's pre-set criteria
Peterseim, M, Papa, C., Wilson, E.,Cheeseman, E., Wolf, B., Davidson, J., Trivedi, R Am. J Ophthalmol. 2014;158:932-938	Compare refractive data and testability of Spot and Plusoptix A09 and compare with traditional cycloplegic retinoscopy	265 children 6.0 +/- 3.4 years	Mean difference in sph, cyl, axis and SE was compared between cyclo. Ret and readings of both devices	Compared with wet ret both devices underplussed and overestimated astigmatism. Spot provided data on a higher % of children but the Plusoptix agreed more closely with cyclo. Ret.	Mean difference used to compare the measurements of Spot and GSE. This method does not examine differences in measurement agreement over a range of measurements

Evan Silverstein, MD; Sean P. Donahue, MD, PhD	2 screening instruments with different referral criteria were evaluated, Spot (sensitive manufacturer's referral criteria) and plusoptiX S08 (specific modified Arthur referral criteria)	2,801 were screened with Spot,307 were referred by Spot and screened by plusoptix, 100 received gold standard examination	All children were screened by the Spot, referred children were then screened using plusoptix. Referred children then received gold standard examination to determine if amblyopia risk factors were present	Plusoptix with modified Arthur referral criteria can be highly specific screening device detecting amblyopia risk factors without missing children with moderate/severe amblyopia	Subjects who passed screening with both devices did not have GSE, assumption that all subjects that were referred by the plusoptix would also be referred by the Spot, small n size
Crescioni et al.	2 devices Spot (V 2.0.16) and Plusoptix S12 (ROC 4.V6.1.4.0) for the detection of astigmatism	209 chidren ages grade 3 to 8	Children screened by Spot had a cycloplegic autorefraction performed by the retinomax and was adjusted by subjective refraction to determine refractive error. ROC analysis of right eye only and sens/spec was calculated	Testability was 97% for the Spot and 54% for the Plusoptix. Sensitivity for detection of astigmatism was 96% for the Spot and 100% for the Plusoptix. Specificity was 87% for Spot and 61% for the Plusoptix. AUC was 0.97 Spot and 0.83 Plusoptix for the detection of astigmatism.	Weaknesses included a relativity small sample size, use of only the right eye for ROC analysis, and older population and not comparing the results of the screening devices with the results of cycloplegic retinoscopy.
Mendez et al.	Spot's V 2.0.16 ability to detect amblyopia risk factors as defined by the 2013 AAPOS vision screening criteria guidelines	219 children average age 60 months	Children were screened by the Spot and then received a GSE with cycloplegic refraction. In 30 of the subjects screening was performed by the ophthalmologists who also performed GSE.	ARF were found in 27 subjects (12.3%) the most common reason being astigmatism. The Spot referred 43 (19.6%) and had a sensitivity of 92.6% specificity of 90.6%. Spot's reason for referral was 31 for astigmatism, 3 for anisometropia, 3 for hyperopia and 3 for myopia.	Weakness of this study include, using only sensitivity and specificity and not ROC for analysis, a relativity small sample size and 30 subjects who were not fully masked for GSE.

Appendix III: Cutoffs for Sensitivity and Specificity at Specificity 0.90

	Senvec	Spvec	SpotSph	SpotCyl	SpotAnisoSph	SpotAnisoCyl
[1,]	0.7355372	0.9005525	1.00	-1.75	1.25	1.50
[2,]	0.7355372	0.9005525	1.00	-1.75	1.25	1.75
[3,]	0.7355372	0.9005525	1.00	-1.75	1.50	1.50
[4,]	0.7355372	0.9005525	1.00	-1.75	1.50	1.75
[5,]	0.7355372	0.9005525	1.00	-1.75	1.75	1.50
[6,]	0.7355372	0.9005525	1.00	-1.75	1.75	1.75
[7,]	0.7107438	0.9005525	1.25	-1.75	0.75	1.00
[8,]	0.7272727	0.9005525	1.75	-1.50	1.25	1.25
[9,]	0.7272727	0.9005525	1.75	-1.50	1.50	1.25
[10,]	0.7272727	0.9005525	1.75	-1.50	1.75	1.25
[11,]	0.6446281	0.9005525	1.75	-2.00	0.75	0.75
[12,]	0.7107438	0.9005525	2.00	-1.50	1.00	1.25
[13,]	0.6776860	0.9005525	2.00	-1.75	1.75	0.75
[14,]	0.7024793	0.9005525	2.25	-1.50	1.00	1.25
[15,]	0.6694215	0.9005525	2.25	-1.75	1.75	0.75
[16,]	0.6859504	0.9005525	2.50	-1.50	1.00	1.25
[17,]	0.6528926	0.9005525	2.50	-1.75	1.75	0.75
[18,]	0.6859504	0.9005525	2.75	-1.50	1.00	1.25
[19,]	0.6528926	0.9005525	2.75	-1.75	1.75	0.75
[20,]	0.6859504	0.9005525	3.00	-1.50	1.00	1.25
[21,]	0.6528926	0.9005525	3.00	-1.75	1.75	0.75

Bibliography

(AAO) American Academy of Opthalmology. Preferred Practice Patterns. Amblyopia. Reviewed 2012.

(AAP) American Academy of Pediatrics. Instrument-Based Pediatric Vision Screening Policy Statement. (2012). Pediatrics; 130;983.

AAPOS 2011. Vision screening laws EDIT

Altman D and Bland J. Diagnostic tests 1: sensitivity and specificity. (June 1994). BMJ. Vol 308. 1552.

Amos JF. Refractive amblyopia Diagnosis and management in vision care. (1987). Boston: Butterworths, 369-408.

(AOA) American Optometric Association. Care of the Patient with Amblyopia. Clinical Practice Guidelines. Reviewed 2004.

Bedell HE, Flom M, Barbeito R. Spatial aberrations and acuity in strabismus and amblyopia. (1985). Invest Ophthalmol Vis Sci; 26:909.

Birch, E., Strauber, S., Beck, R.,, Holmes, J., (2009) Comparison of the Amblyopia Treatment Study HOTV and Electronic-Early Treatment of Diabetic Retinopathy Study Visual Acuity Protocols in Amblyopic Children Aged 5 to 11 years. J AAPOS;13(1): 75-78.

Bland JM, Altman DG. (1999). Measuring agreement in method comparison studies. Statistical Methods in Medical Research. **8** (2): 135–60.

Bland JM, Altman DG (1986). Statistical Methods for Assessing Agreement between two Methods of Clinical Measurement. Lancet **327** (8476): 307–10.

Chua, B and Mitchell, P. 2004 Consequences of amblyopia on education, occupation and long term vision loss. British Journal Opthalmology 88: 1119-1121.

Cotter, S., Cyert, L., Miller, J., Quinn, G., for the National Expert Panel to the National Center for Children's Vision and Eye Health. Vision Screening for Children 36 to <72 Months: Recommended Practices. Optometry and Vision Science. (January 2015). Volume 92. No. 1.

Crescioni, M., Miller, J, Harvey, E. (2015). Accuracy of the Spot and Plusoptix photoscreeners for detection of astigmatism. Journal of AAPOS. Vol. 19 No. 5

Donahue, S., Arnold, R., Ruben, J., Silbert, D., Arthur, B., Neely, D., (Feb. 2012) Revised Guidelines for Reporting Results from Studies of Preschool Vision Screening. JAAPOS. Volume 16 Issue 1.

Donahue, S., Arthur B., Neely D., Arnold R., Silbert D, Ruben J., POS Vision Screening Committee. (2013). Guidelines for Automated Preschool Vision Screening: a 10 year evidence based update. J AAPOS. 17(1):4-8.

FDA) U.S. Food and Drug Administration website. (accessed June 15, 2015). (http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/Overview/GeneralandSpecialControls/ucm055910.htm.

Garry, G., Donahue, S., (2014). Validation of Spot Screening Device for Amblyopia Risk Factors. Journal of AAPOS:18:476-480.

GraphPad Statistics Guide. Interpreting results: Bland-Altman. http://graphpad.com/guides/prism/6/statistics/index.htm?bland-altman_results.htm. Accessed 10/20/15.

Hambidge SJ, Emsermann CB, Federico S, Steiner JF.(2007). Disparities in pediatric preventive care in the United States, 1993-2002. Arch. Pediatr Adolesc Med.;161:30Y6.

Harrad, R., Sengpiel, F., Blakemore, C., (1996). Physiology of suppression in strabismic amblyopia. British Journal of Ophthalmology. 80: 373-377.

Hillis, A., Flynn, JT., Hawkins, BS. (1983). The Evolving Concept of Amblyopia: A Challenge to Epidemiologists. Am J Epidemiology. 118 (192). 205.

(ICO) International Council of Ophthalmology. (Accessed May 29, 2015) Number of Ophthalmologists in Practice and Training Worldwide. http://www.icoph.org/ophthalmologists-worldwide.html.

Jamali P, Fotouhi A, Hashemi H, Younesian M, Jafari A. (2009). Refractive Errors and Amblyopia in Children Entering School Shahrood, Iran. Optom Vis. Sci. Apr. 86(4). 364-9.

Jones, RK, Lee, DN. (1981). Why two eyes are better than one: the two views on binocular vision. J Exp Psychol Hum Percept Perform; 7: 30-40.

Joy S, Davis H, Buckley D. (2001). Is stereopsis linked to hand-eye coordination? Br Orthopt. J;58:38-41.

Kemper AR, Wallace DK, Patel N, Crews JE. (2011). Preschool vision testing by health providers in the United States: findings from the 2006-2007 Medical Expenditure Panel Survey. J AAPOS;15:480Y3.

Klein, R., Wang, Q., Klein, BE., Moss, SE., Meuer, SM. (1995). The relationship of age-related maculopathy, cataract and glaucoma to visual acuity. Investigative Opthalmology Vision Science. 36 (1). 182-191.

Kulp, M., Ciner, E., Maguire, M., Moore, B., Pentimonti, J., Pistilli, M., Cyert, L., Candy, R., Quinn, G., Ying, G. Uncorrected Hyperopia and Preschool Early Literacy. Opthalmology. Article in Press. EDIT

Kulp, M., Ying, G., Huang, J., Maguire, M., Quinne, G., Ciner, E., Cyert, L., Orel-Bixler, D., Moore, B., (2014). Accuracy of Non-Cycloplegic retinoscopy, Retinomax Autorefractor, and SureSight vision screener for detecting significant refractive errors. Invest. Opthalmol. Vis. Sci. 55(3):1378-85.

Lalken A and McCluskey A. (2008). Clinical tests: sensitivity and specificity. Continuing education in anesthesia, Critical Care & Pain. Vol 8 no 6.

Lebow, K. and Campbell, C. (2014). A Comparision of a Traditional and Wavefront Autorefraction. Optometry and Vision Science. Vol. 91. No. 10.

Lamoureux, E, Saw, S., Thumboo, J., Wee, H., Aung, T., Mitchel, P., Wong, T., (2009) The Impact of Corrected and Uncorrected Refractive Error on Visual Functioning: The Singapore Malay Eye Study. Investigative Opthalmology & Visual Science. Vol. 50, 2614-2620.

Levinson DR. (2010). Most Medicaid Children in Nine States Are Not Receiving All Required Preventive Screening Services,OEI-05-08-00520. Washington, DC: United States Department of Health and Human Services, Office of Inspector General; 2010. 2, 4. Available at: https://oig.hhs.gov/oei/reports/oei-05-08-00520.pdf. Last accessed May 2, 2013.

Marsh-Tootle WL, Wall TC, Tootle JS, Person SD, Kristofco RE. (2008). Quantitative pediatric vision screening in primary care settings in Alabama. Optom Vis. Sci.;85:849Y56.

Martinez, V. (2002). Children's Vision Awareness Study: Quanttitative Research Results. Vision Service Plan. Sacremento, CA.

Membreno, J., Brown, M., Brown, G., Sharma, S., Beauchamp, G., (2002). A Cost-utility Analysis of Therapy for Amblyopia. Opthalmology: V 109: N 12.

Mendez, M., Arguello, L., Martinez, J., Vargas, M., Rodriguez, A., Papa, C., Peterseim, C., Viquez, M., Trivedi, R., Peterseim, M. (2015). Evaluation of the Spot Vision Screener in Young Children in Costa Rica. JAAPOS. Vol. 19 No 5.

(MEPEDS) Multi-Ethnic Pediatric Eye Disease Study Group. (2010). Prevalence of myopia and hyperopia in 6- to 72-month-old African American and Hispanic children: the multiethnic pediatric eye disease study. Ophthalmology: 117(1):140-147.

Miller, JM., Dobson, V., Harvey, E.M, Sherrill, DL. (2001) Comparison of preschool vision screening methods in a population with a high prevalence of astigmatism. Investigative Ophthalmology and Visual Science. Vol. 42. Issue 5. Pages 917-924.

Mirzajani, A., Heirani, M., Jafarzadehpur, E., Haghani, H., (2013). A Comparision of the Plusoptix S08 Photorefractor to Retinoscopy and Cycloretinoscopy. Clinical and Experimental Optometry; 96: 394-399.

Modest, J., Johnston, S., Majzoub, K., Moore, B., Trudell, E., Ramsey, J., Vernacchio, L., (2016) Can Chart-Based Vision Screening for Preschool Age Children be Optimized?: Results of a Primary Care Based Quality Improvement Project. Under Review

Murdoch JR, McGhee CN, Glover V. (1991). The relationship between stereopsis and fine manual dexterity: pilot study of a new instrument. Eye; 5: 642-643.

Peterseim, M., Papa, C., Wilson, E., Cheeseman, E., Wolf, B., Davidson, J., Trivedi, R., (2014). Photoscreeners in the Pediatric Eye Office: Compared Testability and Refractions on High Risk Children. Am J. Ophthalmol;158:932-938.

Polling, JR., Loudon SE, Klaver CC. (2012). Prevalence of amblyopia and refractive errors in an unscreened population of children. Optom Vis. Sci. 89 (11) 44-9.
Powell, C. Hatt SR. (2009). Vision screening Programmes for Amblyopia (Lazy Eye). Cochrane Review.

Rahi J, Logan S, Timms C, Russell-Eggitt I, Taylor D. (2002).Risk, causes, and outcomes of visual impairment after loss of vision in the non-amblyopic eye: a population-based study.Lancet. Aug 24;360(9333):597-602.

Ransbarger, K., Dunbar, J., Choi, S., Khazaeni, L., (2013). Results of a Community Vision Screening Program using the Spot Photoscreener. Journal of AAPOS;17:516-520.

Rein, D., Wittenborn, J., Zhang, X., Song, M., Saaddine, J., (2012). Journal Pediatric Opththalmology Strabismus; 49: 146-155.

Satiani, N., and Mutti, D., (2011). Screening for Hyperopia in Infants Using the PowerRefractor. Optometry and Vision Science. Vol. 88. No. 8. PP 095-911.

Schaeffel (1993) inter-individual variability. J Physiol. Feb;461:301-20.

Schimitzek, T., Haase, W., (2002). Efficiency of a video-refractometer used as a screening device for amblyogenic factors. Graefe's Arch. Clin. 240:710-716.

Schimitzek, T., Lagreze, W., (2005). Accuracy of a New Photorefractometer in Young and Adult Patients. Graefe's Arch. Clin. Exp. Ophthalmology. 243: 637-645.

Schmidt P, Maguire M, Dobson V, Quinne G, Ciner E, Cyert L, Kulp MT, Moore B, Orel-Bixler D, Redford M, Ying GS. (2004) Comparison of Preschool Vision Screening Tests as Administered by Licensed Eye Care Professionals in the Vision in Preschoolers Study. Vision In Preschoolers Study Group. Opthalmology;111:637-50.

Shaw JS, Wasserman RC, Barry S, Delaney T, Duncan P, Davis W, Berry P. (2006). Statewide quality improvement outreach improves preventive services for young children. Pediatrics;118:1039Y47.

Shi, Li. (2011) Quantitative Binocular Assessment Using Infrared Video Screening. Dissertation. University of Tennessee Knoxville.

Silbert, D., and Matta, N., (2014). Performance of the Spot Vision Screener for the Detection of Amblyopia Risk Factors in Children. Journal of AAPOS; 18:169-172.

Silverstein, E. and Donahue S. (2014). Field Evaluation of the Spot and PlusoptiX S08 Vision Screeners. JAAPOS Vol. 18 issue 4.

Simons K. (2005).Amblyopia Characterization, Treatment and Prophylaxis. Surv Opthalmol;50:123-66.

Stange KC, Flocke SA, Goodwin MA, Kelly RB, Zyzanski SJ. (2000). Direct observation of rates of preventive service delivery in community family practice. Prev Med;31:167Y76.

Steinman, S., Steinman, B., Garzia, R. (2000). Foundations of Binocular Vision A Clinical Perspective. EDIT publisher

(USCB) United States Census Bureau. (Accessed Sept. 9, 2015). Children Characteristics 2009-2013 American Community Survey 5 year estimates. S0901. American Fact Finder.

(US Dept. Labor) United States Department of Labor Bureau of Labor Statistics. Occupational Employment and Wages (May 2014). 29-1041 Optometrists. Accessed May 29, 2015. http://www.bls.gov/oes/current/oes291041.htm.

(USPSTF). US Preventive Services Task Force. (2011). Vision Screening in Children Ages 1 to 5 years: Clinical Summary of U.S. Preventive Services Task Force Recommendation. Pediatrics;127:340-346.

Van Leeuwen, R., Eijkemans, M.,Vingerling, J., Hofman, A., Jong, P., Simonsz, H., (2007). Risk of Bilateral Visual Impairment in Individuals with Amblyopia: the Rotterdam Study. British J. Opthalmology: 91: 1450-1451

(VIP-HIP study group) Kulp, M., Ciner, E., Maguire, M., Moore, B., Pentimonti, J., Pistilli, M., Cyert, L., Candy, R., Quinn, G., Ying, G., (2016). Uncorrected Hyperopia and Preschool Early Literacy. American Academy of Opthalmology. Article in Press.

(VIP) The Vision in Preschoolers Study Group. (2004). Comparison of Preschool Vision Screening Tests as Administered by Licensed Eye Care Professionals in the Vision in Preschoolers Study. Ophthalmology. 111(4). 637- 650.

(VIP) The Vision in Preschoolers Study Group. (2005). Preschool Vision Screening Tests Administered by Nurse Screeners Compared with Lay Screeners in the Vision in Preschoolers Study. Investigative Opthalmology and Visual Science. Vol 46. No 8.

(VIP) The Vision in Preschoolers Study Group. (2007). Children Unable to Perform Screening Tests in Vision in Preschoolers Study: Proportion with Ocular Conditions and Impact on Measures of Test Accuracy. Investigative Ophthalmology and Vision Science; 48-83-7.

Virgilli, G., Angi, M., Heede, S., Rodriguez, D., Botegga, E., Molinari, A., (2007). PowerRefractor vs. Cannon R-50 Autorefraction to assess refractive error in children: a community based study in Ecuador. Optom Vis Sci.;84(2):144-8.

Von Noorden, G., Campos, E. (2002). Binocular Vision and Ocular Motility Theory and Management of Stabismus.

Welch Allyn Communications. June 15, 2015. How Spot Works.

Wen G, Tarczy-Hornoch K, McKean-Cowdin R, et al. (2013). Prevalence of myopia, hyperopia, and astigmatism in non-Hispanic white and Asian children: multiethnic pediatric eye disease study. Ophthalmology: 120(10):2109-2116.

WHO. (2001). Model Regulatory program for medical devices International Guide. Pages 21-23.

Wilson JMG, Jungner G. (1968) Principles and practice of screening for disease. WHO Chronicle Geneva:World Health Organization. 22(11):473. Public Health Papers, #34.

Yajun M., Hua B., Edgar E., Gang D., Nan W., Ning H., Xuehan Q., Xiaorong L., (2016). Performance of Spot Photoscreener in detecting amblyopia risk factors in Chinese Pre-schoolers and school aged children attending an eye clinic. PLoS ONE 11(2): e0149561. doi:10.1371/journal.pone.0149561.